THE PHYSICAL SOCIETY

Reports on
PROGRESS IN PHYSICS

Price to non-Fellows 12*s.* 6*d. net*

Published by
THE PHYSICAL SOCIETY
1 Lowther Gardens, Exhibition Road
London, S.W. 7

Printed at
THE UNIVERSITY PRESS, CAMBRIDGE
1934

THE
PHYSICAL
SOCIETY

PRINTED IN GREAT BRITAIN

CONTENTS

PAGE

Quantum and wave mechanics. By Professor G. TEMPLE 1

The spiral nebulae and the expansion of the universe. By G. C. McVITTIE 24

Classical physics: surface tension. By ALLAN FERGUSON 30

Burgers's theory of turbulence. By G. F. P. TRUBRIDGE 43

Sound, covering the years 1932 and 1933. By E. G. RICHARDSON . . 63

Spectroscopy 102

§ 1. Band spectra. By R. W. B. PEARSE
§ 2. Infra-red photography. By H. DINGLE
§ 3. Grating theory. By H. DINGLE
§ 4. Ultra-violet spectroscopy and term-analysis of spectra. By H. DINGLE
§ 5. Microphotometry. By H. DINGLE
§ 6. Absorption and fluorescence spectra of crystals. By H. DINGLE
§ 7. The hyperfine structure of spectrum lines. By E. GWYNNE JONES
§ 8. Historical notes on wave-length standards. By C. V. JACKSON
§ 9. More recent progress in wave-length standards. By C. V. JACKSON
§ 10. Stellar research. By A. HUNTER
§ 11. Solar research. By A. HUNTER
§ 12. Instrumental technique in astrophysical spectroscopy. By A. HUNTER

Optics. By MEMBERS OF THE STAFF OF THE NATIONAL PHYSICAL LABORATORY 131

§ 1. Optical instruments. By T. SMITH
§ 2. Refractometry. By J. S. ANDERSON
§ 3. Vision. By J. GUILD
§ 4. Colorimetry. By J. GUILD
§ 5. Photometry. By H. F. BUCKLEY

Heat. By J. H. AWBERY 161

Electrical phenomena at extremely low temperatures. By Professor J. C. McLENNAN 198

PAGE

X rays. By THE X-RAY STAFF, PHYSICS DEPARTMENT, NATIONAL PHYSICAL LABORATORY (*General Editor*: G. W. C. KAYE) 228

§ 1. Production of X rays. By G. E. BELL

§ 2. Ionization in gases. By I. BACKHURST

§ 3. Absorption of X rays. By I. BACKHURST

§ 4. Scattering of X rays. By I. BACKHURST

§ 5. Measurement of intensity, quality, etc. By W. BINKS and G. E. BELL

§ 6. X-ray spectra. By G. SHEARER. With a sub-section on continuous spectrum. By G. E. BELL

§ 7. X-ray optics. By G. SHEARER

§ 8. Applications of X rays. By W. A. WOOD

The Physics of the Atom. By Professor E. N. DA C. ANDRADE . . . 269

Electric and magnetic measurements. By L. HARTSHORN 321

QUANTUM AND WAVE MECHANICS

By Professor G. TEMPLE, Ph.D., D.Sc., F.Inst.P.
University of London, King's College

§ 1. INTRODUCTION

Although this article is officially the first of a series of annual reports on contemporary progress in quantum and wave mechanics, it seems to be advisable to preface the report of recent advances with a general survey of the whole subject. This appears to be especially desirable because the branch of physics discussed in this article has developed along lines which diverge widely from those branches of classical physics which are discussed in the other sections of the series. It is therefore natural to begin with a general account of the nature of the quantum theory indicating what affinities it has with other departments of mathematical physics.

After this introductory account of the nature of modern quantum theory, we shall sketch briefly the accepted principles of the theory. These fall naturally into two classes—first, the general principles which are applicable to any physical system, and, secondly, the special principles which are characteristic of particular systems. The historical development of both groups of principles falls into two periods which may be roughly described as the period of discovery and the period of interpretation. The period of discovery includes the pioneer investigations of Heisenberg, de Broglie and Schrödinger which issued in the formal statement of the rules of calculation in quantum theory. The period of interpretation refers to the subsequent researches of Dirac, v. Neumann and others which had as their object the elucidation of the physical significance of these formal rules.

With this historical survey as a background, we shall then proceed to summarize the novel and important developments made by Eddington during the last two or three years. These investigations lie in the border-land (or no-man's-land!) between the theory of relativity and the quantum theory, and the object of these investigations is to effect for microphysics what the theory of relativity has effected for macrophysics, i.e. to carry out a geometrization of the quantum theory which shall enable the characteristic properties of the microphysical elements to appear as natural and inevitable consequences of the quasi-geometrical structure of the quantum world. These investigations lead to important conclusions regarding the properties of the electron, proton and the positron. In a discussion of these physical elements it is convenient to include also the neutron (whose elementary character is still a matter of debate), and the universe itself, considered as one gigantic atom.

The quantum theory of the universe, considered as a system interacting with an electron, connects up the researches summarized in this article with the investigations summarized by Dr McVittie in the article on the spiral nebulae and the expansion of the universe.

In view of the wealth of literature dealing with the quantum theory which has been published during the last two or three years, it has been a matter of serious embarrassment to make an adequate selection for the purposes of this article. The indiscriminate inclusion of all the researches that have been published would be futile and impossible, and a mere catalogue of results would be useless. Under these circumstances it appeared to be best to limit this article severely to the topics enumerated above, i.e. to a review of the developments which have been made in the principles of the subject, and in its applications to the fundamental physical elements.

§ 2. THE NATURE OF THE QUANTUM THEORY

An air of mystery has always hung over the quantum theory from its first appearance in the world of physics. The mystery arises from the fact that in the quantum theory it appears at first sight that conclusions about the physical world have been deduced from principles which are mere mathematical definitions and conventions. A cursory perusal of any of the standard textbooks on the quantum theory will show that the theory is developed primarily as a branch of pure mathematics, of which only a few of the terms and relations receive a physical interpretation, and that only at a comparatively late stage in the development of the theory.

To dissipate this undoubted mystery in which the subject has developed, it would, of course, be necessary to develop the subject in such a way that the fundamental assumptions should emerge as the expressions of some physical principles susceptible of a direct experimental verification, but so far this ideal has not been satisfactorily realized by any of the exponents of the theory. At the present stage of development of the theory, the most satisfactory mode of treatment is, therefore, one in which the present paradoxical character of the theory is sharply emphasized by contrasting the purely mathematical scheme of the theory with its rather tentative physical interpretation. This aspect of the theory can perhaps be most readily appreciated by making a general comparison between quantum theory and statistical mechanics.

The latter subject is also surrounded by the same air of mystery as that which still envelops the quantum theory. The mathematical scheme which forms, as it were, the skeleton of statistical mechanics has been a subject of remarkable development and has attained a high degree of technical perfection. But at the same time it must unfortunately be admitted that the physical significance of the methods used and the results obtained in the mathematical scheme are still most obscure. There is, however, a remarkable resemblance in detail as well as in general between the quantum theory and statistical mechanics, and in this section it is proposed to make a detailed comparison of these two branches of mathematical physics the basis of our study of the nature of the quantum theory.

The quantum theory and statistical mechanics may be compared both as regards the subject-matter with which they deal and as regards the general character of the results which they predict, and they may both be contrasted in these same two respects with classical dynamics. Classical dynamics purports to discuss properties of individual systems and to enable us to calculate the actual values of the physical variables which are characteristic of these systems. But both in quantum theory and in statistical mechanics we are not concerned with individual systems, but with assemblies, aggregates or *ensembles* of systems, and the results obtained are not the actual values of variables but are mean values obtained by a process of averaging over all the systems comprising an assembly. The main difference between these two subjects is that in statistical mechanics the systems considered are large composite systems, such as a gas, consisting of many smaller sub-systems; while in quantum mechanics, at its present stage of development, the systems considered are the ultimate or penultimate elements of the physical world, such as individual electrons, atoms and molecules. This difference between the two subjects is one of degree rather than kind, and there is no doubt that subsequent developments in the quantum theory will invade the domain of statistical mechanics. The first steps in this direction have already been taken by v. Neumann[1].

The assemblies or aggregates which are considered in the quantum theory and in statistical mechanics are not real collections of systems actually existing in the external world. They are merely convenient logical fictions introduced into the mathematical scheme to assist in the calculation of numerical values, which, surprisingly enough, do appear to have reference to the individual systems actually existing in the physical world. The simplest example of an assembly studied from statistical mechanics is the microcanonical ensemble. The composite systems which form this assembly are specified as containing a definite number, N, of similar sub-systems occupying a definite volume, V, and possessing a definite amount of total energy, E. The composite systems are chosen so that each system in the assembly represents one of the possible methods by which total energy can be distributed among the sub-systems. Thus each possible method of distribution is represented exactly once in the whole assembly.

The fundamental and characteristic assumption of statistical mechanics is that the actual distribution of energy among the sub-systems in any actual composite system studied in the laboratory may be determined by taking the average of all the possible distributions represented in a complete assembly. Similarly, it is assumed that the actual value of any variable, such as the pressure, characterizing any actual composite system, such as a gas, may be calculated by the same process of averaging from the values of pressure to be contributed to each of the composite systems in the fictitious assembly.

The same general method for the calculation of average or effective values is adopted in the quantum theory. Here again, the theory does not contemplate individual systems, but deals only with certain canonical assemblies, which are of course convenient logical fictions, introduced for the purpose of computation. Just as in statistical mechanics these assemblies are not determined by a complete specifica-

tion of all the characteristics of the systems which they comprise. The standard method of specifying a quantum assembly is as follows: Each of the individual systems of which it is composed is defined to have the same energy, or possibly the same set of components of momentum. Of course, this leaves the values of the co-ordinates of the individual systems entirely undetermined. The specification of the assembly is therefore completed by specifying the distribution of the values of the co-ordinates among the individual systems of the assembly. When the components of momentum have been specified the standard distribution is that in which equal numbers of systems are found in congruent regions of the co-ordinate space. For example, an assembly of elementary structureless particles is specified by prescribing the three components of the momentum common to each particle, say p_x, p_y, p_z, and by prescribing the distribution of the values of the spatial co-ordinates x, y and z. In this example the number of systems in any region of the x, y, z space is taken to be proportional to the volume of that region.

The specification of the values of the momenta is not the only way of defining a quantum assembly. Such an assembly can also be defined by specifying the values of the co-ordinates common to all the systems of the assembly or by a mixed scheme of definition in which some of the defining variables are momenta and some are co-ordinates. In each case the definition is completed by specifying a standard distribution of the unspecified co-ordinates of momenta among the individual systems.

p_x, p_y, p_z
m_x, m_y, m_z

The calculation of average values in the quantum theory also parallels the similar calculation in statistical mechanics. For example, in the case of an assembly in which the defining variables are the three components of momenta, p_x, p_y, p_z, the average or effective value of the components of angular momentum, m_x, m_y, m_z, is obtained by averaging over all the systems composing the complete assembly. In the application of the theory to spectroscopy, the assemblies considered are those in which the systems are atoms or molecules each with the same energy. The effective values of the atomic momenta, magnetic moment or electric moment, are obtained by averaging over all the systems in the assembly. In this case the variable determining the distribution of the systems is the variable which is dynamically conjugate to the energy, viz. the time. In this case, therefore, the average over all the systems of the assembly is equivalent to a long time average taken over an individual system.

Both in statistical mechanics and in quantum theory it is assumed that the average values calculated by the methods outlined above can be identified with the values which are obtained in actual experiments. Neither in the case of statistical mechanics nor in the case of quantum theory is there any really compelling argument for this identification. Both branches of physics exhibit the same two characteristics of an elaborate mathematical technique which is only brought into contact with reality by the unjustified identification of the final results of the calculation with various physical entities. This is a real problem in both theories—to give a physical interpretation to the whole of the mathematical scheme and not merely to its end products. A brief sketch of the principal attempts to solve this problem is given in the next section.

§ 3. THE GENERAL PRINCIPLES OF THE QUANTUM THEORY

The comparison of the quantum theory and of statistical mechanics which we have made in the previous section must be supplemented by an account of the vital distinction between these two theories. The fundamental divergence between the two theories arises from their attitude towards the significance of the values of the variables which are characteristic of physical systems.

In statistical mechanics, as in the whole of classical physics, it is assumed that these numbers have an intrinsic value as representing completely the actual state of the systems considered. But in quantum theory it is maintained that these numbers are not primarily characteristic of the systems to which they refer, but of the responses of these systems to certain experimental stimuli. This emphasis in quantum theory on what may be called for brevity the biological significance of the variables, is of fundamental importance. The distinction between the two points of view becomes clear when we consider its influence on the calculation of average values in these two branches of physics.

In statistical mechanics, although we are not able to calculate explicitly the actual characteristics of the individual systems of an assembly, nevertheless we implicitly assume that each system is characterized by its full complement of variables which serve to satisfy it completely. In quantum theory, however, this assumption is challenged. The basis of the theory is in fact the self-denying ordinance that no assumption of this kind is to be admitted until it has been justified experimentally. Thus, in quantum theory, to assert that in a certain system certain variables have certain specified values means that it is physically possible to perform experiments on the system which will actually yield the specified values. This apparently innocuous assumption becomes of dominant importance in the theory as soon as it is realized that in the domain of microphysical elements the mere "observation" of a system may produce a change in the state of that system.

The argument relating to the nature of microphysical measurement and its influence on quantum theory is due to Heisenberg[2] and runs as follows: The outstanding inference to be drawn from microphysical experiments, i.e. experiments on the ultimate elements of matter, is that the interaction of these elements with one another is always subject to a certain "quantization." Thus the emission and absorption of energy in the form of radiation of frequency ν always take place in units of magnitude $h\nu$, where h is Planck's constant. Now, in any microphysical measurement the action of the system to be measured upon the apparatus for observation, by which the system produces an observable record, must be accompanied by some reaction of the apparatus upon the system. This, of course, is true both of microphysics and of macrophysics, but in macrophysics, i.e. physics of matter in bulk, it is possible to make this reaction negligibly small in comparison with the values of the quantities to be measured. But in microphysics the quantization of the processes of interaction between the ultimate elements of matter and of radiation makes it impossible to reduce the amount of interaction to a negligible quantity. We must, therefore, be prepared to find that in microphysics any process of measurement will

ν

h

usually produce some change in the system to be measured. Under these circumstances, the results of an experiment can no longer be regarded as entirely characteristic of the system measured, but must rather be regarded as characterizing a particular kind of interaction between the system and the measuring apparatus.

This disturbing conclusion can only be avoided if we limit ourselves to those special experiments which do not produce any change in the properties of the system measured[3]. Such experiments are obviously characterized by the fact that when repeated any number of times they yield the same results as they did on the first occasion. This elementary statement can be made to serve as a basis of the mathematical technique of the quantum theory[4].

The special type of experiment whose result is unchanged by a repetition of the experiment may be conveniently described as a "simple observation." In order to express the properties of a simple observation in symbolic form, it must be recognized that the values of the different physical characteristics may be continuously distributed over a wide range, as in the case of the kinetic energies of atoms and molecules, or they may be confined to certain discrete values as in the case of the components of angular momentum. The general theory must take into account both of these possibilities. Thus, the simplest general type of observation is that the value of a variable y lies within a certain range $y_n \leqslant y < y_{n+1}$, where the y_n's are convenient numbers dividing up the whole possible range of values of the variable y. A simple observation of this type will be denoted by the symbol Y_n. The set of symbols $Y_1, Y_2, \ldots$ represents those simple observations which show that for the system observed the variable y lies within the corresponding ranges $y_1 \leqslant y < y_2, y_2 \leqslant y < y_3, \ldots$. The carrying out of a simple observation is clearly analogous to the process of spectral analysis. If the process represented by the symbol Y_n is carried out twice, this double process will naturally be represented by the symbols $Y_n Y_n$, or Y_n^2. Moreover, since this process is a simple observation, the process represented by Y_n and Y_n^2 must be equivalent. Hence we must write

y_n, y

Y_n

$$Y_n^2 = Y_n.$$

This operational equation does not imply of course that Y_n is equal to one of the numbers unity or zero, for it will be clear on reflection that this conclusion is only justified when Y_n possesses a reciprocal, and it is obvious that the process of simple observation is necessarily irreversible and therefore that no reciprocal process can exist. The simplest example of a mathematical process which satisfies this operational equation is the process of projecting any arbitrary vector in a specified direction. The repetition of such a process of projection (of course in the same specified direction) does not disturb the result of the first application of the process. Hence if Y_n is a projective operator, we have that $Y_n^2 = Y_n$. There is thus nothing intrinsically absurd about such an equation, and it is clearly possible to construct mathematical operators which shall enjoy the properties possessed by the symbol Y_n.

In order that the complete process of measurement represented by the symbols $Y_1, Y_2, \ldots$ may be physically significant, it must satisfy the three conditions of consistency, exclusiveness and completeness. The condition of consistency is that

which we have just obtained, and is represented by the equation $Y_n^2 = Y_n$. The condition of exclusiveness is that application of the simple observation represented by Y_k to the results of the observation represented by Y_n ($k \neq n$) shall yield a null result, for it will be impossible to obtain systems for which the variable y shall lie in the two exclusive ranges

$$y_k \leqslant y < y_{k+1} \quad \text{and} \quad y_n \leqslant y < y_{n+1}.$$

This property of the Y_n's is symbolized by the equation

$$Y_k Y_j = 0, \quad \text{if } k \neq j.$$

Finally, the condition of completeness is that every system in a given assembly must be obtained by application of one of the processes of observation represented by $Y_1, Y_2, \ldots$. This property of the Y_n's is represented by the equation

$$Y_1 + Y_2 + \ldots = 1.$$

It has already been remarked that projective operators satisfy the equation $P^2 = P$, and it is clear that a set of symbols satisfying the three conditions obtained in the last paragraph can be represented by a set of projective operators. If the number of operators is n, it is necessary to work in n-dimensional space. Choosing n orthogonal direction in this space, we can represent the symbols $Y_1, Y_2, \ldots$ by the n operations of projection in these n directions. These n projective operators evidently satisfy the three fundamental equations. This completes all that need be said concerning the representation of the simple observation of a single variable Y.

When, however, we proceed to consider the measurement of two variables, say Y and Z, we shall have to consider two sets of simple observations $Y_1, Y_2, \ldots$ and $Z_1, Z_2, \ldots$, together with the associated sets of projective operators. It is then necessary to consider the relation between the two sets of n orthogonal directions which are respectively associated with the two sets of projective operators. This relation is expressed in terms of the concept of transition probabilities.

Each of the n directions associated with the set of symbols $Y_1, Y_2, \ldots$ represents an assembly of systems for which the variable Y lies in a certain definite range, say

$$y_n \leqslant y < y_{n+1}.$$

If the set of observations $Z_1, Z_2, \ldots$ are applied to the systems in such an assembly, certain of the systems will respond to the simple observation Z_1, and after the application of this observation will have values of z in the range $z_1 \leqslant z < z_2$. Certain of the systems will respond to the observation Z_2, and will subsequently have corresponding values of z, etc. Thus the physical relation between the two sets of simple observations $Y_1, Y_2, \ldots$ and $Z_1, Z_2, \ldots$ is given by the probabilities that a system in the state determined by Y_k will pass over into the state determined by Z_j, and that a system in the state determined by Z_j shall pass over into the state determined by Y_k. Assuming that these two transition probabilities, i.e. from Y_k to Z_j and from Z_j to Y_k, are equal, we can easily deduce what is the corresponding mathematical relation between the two sets of n directions associated with the two sets of simple observations. The mathematical relation is that the transition probability connecting the states specified by Y_k and Z_j is equal to the square of the cosine

of the angle between the two directions associated with these two symbols. There is no direct evidence for the assumption that the value of transition probabilities is independent of the direction of the transition, but this assumption appears to be a natural extension of the principle of detailed balancing as used in statistical mechanics, and it is clearly the simplest assumption of this kind which can be made. The resulting geometrical picture of the relations between the two sets of n fundamental directions is extremely simple and allows us to develop a method of representing physical variables by geometrical operators of a more general kind than the projective operators introduced above.

$p(x)$ Before proceeding to develop this theory, it is convenient to define first the "spectrum" of a given variable y. Let $p(x)$ denote the fraction of the total number of systems in a given assembly for which the value of the variable y is less than the number x. Then $p(x)$ is a steadily increasing function of x which may or may not be continuous. The points at which it is discontinuous form the discrete spectrum of
a the variable y; the discontinuity in $p(x)$ at a point a measures the fraction of systems in the given assembly for which $y=a$. If $p(x)$ is continuous at an assigned point, it must possess a derivative at that point and this derivative may be positive or zero. If it is positive the point is said to belong to the continuous spectrum y: the remaining points at which $p(x)$ is continuous and possesses a zero derivative do not form part of the spectrum of y. The spectrum of a variable may in general be partly continuous and partly discrete, but it will considerably simplify the subsequent development of this section if we limit ourselves to the case of a variable y with a discrete spectrum, $y_1, y_2, \ldots$. This limitation considerably simplifies the analysis and the results obtained are easily extended to the more general case by the replacement of certain sums by definite integrals.

To give an adequate discussion of the theory of the representation of variables by linear operators, it would be necessary to commence with a summary of the general theory of linear operators, vectors and matrices. This purely mathematical investigation can, however, be short-circuited if we adopt the device of regarding the symbols such as $Y_1, Y_2, \ldots$ and $Z_1, Z_2, \ldots$, which represent simple observations, as hypercomplex numbers which obey all the laws of ordinary algebra except the commutative law of multiplication.

The necessity for suspending the commutative law of multiplication is evident when we recall that the two operations represented by the products, YZ and ZY, represent two distinct physical processes which will in general leave the systems upon which they are executed in different states. Since, however, these symbols represent operations, it is clear that they will obey the associative law of multiplication, i.e.

$$X(YZ)=(XY)Z.$$

We shall now show that in a certain conventional sense the variable y is represented by the operator Y, which is defined by the following equation:

$$Y=y_1Y_1+y_2Y_2+\ldots,$$

i.e. Y is the sum of a number of terms each of which consists of the product of one

of the numbers in the discrete spectrum of Y by the corresponding projective operator.

To establish this fact we note first that the n orthogonal directions which are respectively associated with the set of projective operators $Y_1, Y_2, \ldots$ can be regarded as furnishing a geometrical representation of n assemblies, each of which is homogeneous with respect to the variable y, and in which the variable y has the definite values $y_1, y_2, \ldots$. The unit vectors drawn from the origin in these n orthogonal directions will be represented by the symbols $\eta_1, \eta_2, \ldots$, which will be called the proper vectors of the variable y or of the operator Y. When the operator Y acts upon one of these vectors, η_k, it yields a result $Y_k\eta_k = y_k\eta_k$, i.e. it yields the original vector unaltered in direction, but increased in magnitude in the ratio y_k to 1. η_n

Secondly, we introduce the following definition and notation for the scalar product of two vectors: The scalar product of two vectors α and β is defined to be the real number which is obtained by multiplying together the magnitudes of α and β and the cosine of the angle between them, and it is represented by the symbol (α, β). When α and β are unit vectors representing two homogeneous assemblies, the square of this scalar product will clearly represent the probability of a system making a transition from one such assembly to the other. (α, β)

To show in what way the operator Y represents the variable y, we consider the scalar product $(\alpha, Y\alpha)$, where α is a unit vector which represents some definite pure state of the system considered. The vector α can be expressed in terms of the vectors $\eta_1, \eta_2, \ldots$, which represent the pure states in which the variable y has the definite values $Y_1, Y_2, \ldots$. If the expression for α is

$$\alpha = c_1\eta_1 + c_2\eta_2 + \ldots,$$

then it will be clear that the coefficients $c_1, c_2, \ldots$ in this expansion are the direction cosines of the vector α referred to the co-ordinate system determined by the vectors $\eta_1, \eta_2, \ldots$. Hence, the squares of these coefficients or, to be accurate, the squares of the moduli of these coefficients are the transition probabilities from the state represented by α to the states represented by $\eta_1, \eta_2, \ldots$. Now, c_n

$$Y\alpha = y_1c_1\eta_1 + y_2c_2\eta_2 + \ldots.$$

Hence

$$(\alpha, Y\alpha) = c_1^2y_1 + c_2^2y_2 + \ldots,$$

since the vectors $\eta_1, \eta_2, \ldots$ are unit orthogonal vectors. The expression on the right-hand side of the last equation is clearly the average or effective value of the variable y in the assembly represented by the vector α. It is in this sense that the variable y is represented by the operator Y, viz. that for any assembly represented by a vector α the effective value of the variable y is always given by the scalar product $(\alpha, Y\alpha)$.

It may be shown that this representation is unique in the sense that Y is the only linear operator for which the preceding result is universally true. It is not difficult to show that if x and y are two variables which are respectively represented by the operators X and Y, then the sum of these two variables, $x+y$, is represented by the sum of the two operators $X+Y$. Similarly, if c is any ordinary real number, it follows that the operator representing the variable cx is cX. These results are true

whether the variables x and y are compatible or incompatible. It may also be shown that the operator representing x^2 is X^2, and similarly for any positive or negative integral power of x. These results are comparatively trivial. The most interesting result is that which relates to the representation of the product xy. Since

$$xy = \tfrac{1}{4}(x+y)^2 - \tfrac{1}{4}(x-y)^2,$$

it follows by combining the results stated above that the product xy is represented not by XY or by YX, but by $\frac{1}{2}(XY+YX)$. Of course in the special case when the variables x and y are compatible, the operators X and Y will commute and the product xy can be represented indifferently by XY or YX. But, in general, the operator which represents xy must be expressed as the mean of XY and YX.

§ 4. THE SPECIAL PRINCIPLES OF THE QUANTUM THEORY

The mathematical scheme developed in the last section from the general principles of the quantum theory requires to be supplemented and adapted for detailed application to special physical systems by the adjunction of a number of special principles. The detailed application of the quantum theory to particular physical systems has been naturally a matter of urgent importance for the testing of the theory by a comparison of these predictions with the results of experiments, and in view of this urgency the creators of the quantum theory have had perforce to be content with certain more or less plausible assumptions which served as temporary expedients and which, it is hoped, will one day be replaced by more profound and permanent principles.

The most important method which has been used for the application of the quantum theory to special systems is an appeal to what is known as the *correspondence principle*. This principle was originally simply a body of some empirical rules for determining under what conditions the methods and concepts of classical physics or quantum physics should be applied to the various problems under discussion. The principle received a more precise interpretation in the hands of Niels Bohr, but his investigations have now been definitely superseded in the modern form of the quantum theory. In these more recent developments the correspondence principle determines the analogy which exists between the entities and relations of classical physics and of quantum physics. It indicates how the variables of classical physics are to be replaced by linear symmetric operators in quantum physics and how the equations of motion, etc. of classical physics are to be modified in the quantum theory.

The method of application of the correspondence principle and the results which are deduced by its aid are studied in more detail later in this section. It must be mentioned at once, however, that it appears that this principle is only effective in dealing with non-relativistic problems of material systems, i.e. it is not applicable to radiation, or to those problems in which the variation of mass with velocity is no longer negligible. In dealing with relativistic problems considerable help has been

obtained from the application of a second general principle—*the transformation theory of relativity*. The relativistic methods introduced by Dirac have been further developed by Eddington in such a way as to suggest the possibility of a geometrization of quantum theory analogous to the geometrization of classical dynamics by the general theory of relativity.

Quite recently a third new and important principle has been introduced into quantum theory by Eddington. This principle, which so far has not received a name, will be referred to in this report as the *principle of physical abstraction*. It is a principle of a much deeper and more metaphysical character than the preceding two. It arises from a consideration of what is involved in the concept of a "system" or an "isolate," and it appears to be related, though perhaps vaguely, to the holism of General Smuts and the Gestalt psychology of Köhler. We shall now proceed to study these principles in more detail.

The correspondence principle. The possibility of the representation of variables by operators has been shown in § 3 to be a consequence of the general characteristics of microphysical measurement. Some writers, however, especially J. v. Neumann[5], have preferred to accept this operational representation as a fundamental axiom of the quantum theory. v. Neumann assumes that any physical variable can be represented by a linear symmetric operator, and, conversely, that any linear symmetric operator represents some physical variable. Furthermore, he assumes that the representation is of such a kind that if the variables x and y are represented by the operators X and Y, then $ax+by$ is represented by $aX+bY$, where a and b are ordinary numbers. Finally, he assumes that any positive integral power of the variable x is represented by the corresponding power of the operator X. It follows at once from these assumptions that, since the product of any two variables x and y can be expressed in the form

a, b

$$xy=\tfrac{1}{4}(x+y)^2-\tfrac{1}{4}(x-y)^2,$$

it will be represented by the operator $\frac{1}{2}(XY+YX)$.

In this scheme the fundamental problem is to determine the average or the effective value of a physical variable x in an assembly of systems α. This effective value is denoted by the symbol $E_\alpha(x)$. In virtue of the axioms laid down in the preceding paragraph, it follows that with any assembly there is associated a "certain statistical" operator A, such that

α
$E_\alpha(x)$
A

$$E_\chi(x)=\chi(AX),$$

where the symbol $\chi(R)$ denotes the trace or characteristic of the operator R, i.e. the sum of the diagonal elements of this operator in its matrix representation. It is also shown by v. Neumann that if the assembly α is pure, or possesses the maximum degree of homogeneity, then the statistical operator A is a projective operator such that $A^2=A$. This important result links his investigations with the results obtained at the end of the last section, for it is easy to show that

$\chi(R)$

$$\chi(AX)=(\alpha, X\alpha),$$

where α is the unit proper vector of the operator A, which serves equally well to specify the assembly of systems which is being studied.

A completely different method of developing the fundamentals of the quantum theory has been devised by Dirac[3]. The essence of Dirac's method is to effect the widest possible generalization of the system of wave mechanics developed by de Broglie and Schrödinger. These writers have familiarized physicists with the idea that with material particles are associated systems of waves, which serve to guide the particles and to restrict their possible motions to those permitted by the quantum theory. The essential feature of any system of wave mechanics is the possibility of superimposing two different systems of waves so as to obtain a third system, which like the component systems shall equally well represent a possible state of some material system. It is this feature of wave mechanics which is made fundamental in Dirac's scheme.

The fundamental assumption made by Dirac is that it is possible to represent any assembly of systems by a vector in a many-dimensional space. The representation by wave functions is a particular case of the general representation by vectors which arise when the number of co-ordinates in the space is continuously infinite. It is assumed that the probability of agreement of any two states represented by unit vectors, α and β, is equal to $|(\alpha, \beta)|^2$, i.e. the square of the modulus of the scalar product of these two vectors. It is now easy to infer the possibility of representing physical variables by operators; for if $\alpha_1, \alpha_2, \ldots$ are the vectors representing the pure states of a variable A with the spectrum of values A_1, A_2, then with this variable we can associate the operator A defined as the sum $a_1A_1+a_2A_2+\ldots$, where $A_1, A_2, \ldots$ are the projective operators for the vectors $\alpha_1, \alpha_2, \ldots$.

The principal results obtained from the application of the correspondence principle are Heisenberg's exchange relations and the form of the equations of motion in the quantum theory. Dirac's deduction of the exchange relations is extremely simple and ingenious. It is well known that in the advanced theory of classical dynamics an important part is played by expressions known as the "Poisson brackets." Dirac assumes that the Poisson brackets $[x, y]$ of any two variables x and y can be represented by an operator (which is temporarily denoted by the symbol $[X, Y]$) in such a way that the operators representing the Poisson brackets enjoy the same properties as the Poisson brackets themselves. It is then possible to show that

$$[X, Y]=c\,(XY-YX),$$

where c is an ordinary number which must be purely imaginary if x and y are real variables. Now if a dynamical system is specified by a set of co-ordinates $q_1, q_2, \ldots$, together with the conjugate momenta $p_1, p_2, \ldots$, the Poisson brackets of any pair of these variables are easily found to have the following values:

$$[q_j, q_k]=0, \quad [p_j, p_k]=0,$$
$$[q_j, p_k]=0 \quad (j\neq k), \quad [p_j, q_j]=1.$$

From these values we can deduce at once the values of the commutators, $XY-YX$, of the operators X and Y representing any pair of variables. We have, obviously, that every pair of operators commutes except conjugate pairs for which

$$P_jQ_j-Q_jP_j=c^{-1}I.$$

If c is identified with i/h, where h is Planck's constant divided by 2π*, these equations are equivalent to Heisenberg's exchange relations. h

The correspondence principle can also be used to obtain the quantum analogue of the classical equations of motion. In classical dynamics the equation of motion of a variable x is simply an explicit expression for the velocity or time derivative of x in terms of the instantaneous value of x and the contemporaneous values of the other variables of the system. In quantum theory the equation of motion of an operator X is an explicit expression for an associate operator $\dot{X}$ called by analogy "the velocity of X." This operator is defined by the condition that the average value of $\dot{X}$ shall be equal to the rate of change in the average value of X, i.e. $\dot{X}$

$$(\alpha, \dot{X}\alpha) = \frac{d}{dt}(\alpha, X\alpha),$$

for any vector α. Now it may be shown that the sequence of changes in the vector α representing any assembly are appropriately represented by an equation of the form

$$\alpha_t = \exp(-itW/h)\,\alpha_0, \qquad i, W$$

or

$$d\alpha_t/dt = -(i/h)\,W\alpha_t,$$

whence it may be deduced that

$$\dot{X} = (i/h)(WX - XW).$$

The last equation indicates the general form of the analogue of the equation of motion for any operator in quantum theory. By comparing this result with the classical equations of motion expressed in Hamiltonian form and by using Heisenberg's exchange relations it may be inferred that W is the operator representing the Hamiltonian function of the system considered. This result allows us to apply the quantum theory to any non-relativistic dynamical system.

The special theory of relativity. In applying quantum theory to problems in which the variation of mass with velocity is not negligible, it is necessary to make use of the *general transformation theory of relativity*. Dirac[3] has obtained some striking results by direct application of relativistic principles to the quantum equation of motion, i.e. the wave equation. We have already seen that the quantum equation of motion involves the Hamiltonian operator W, which represents the energy linearly. If this equation is to be relativistically invariant, it must also involve the momentum operators in a linear manner, i.e. the Hamiltonian operator W must not be expressed as a quadratic function of the momentum operators p_x, p_y, p_z, but as a linear function of these operators. The exact form of the Hamiltonian operator W is fixed by the further consideration that the equation of motion must be equivalent to the classical relativistic relation

$$w^2 = p_x^2 + p_y^2 + p_z^2 + m_0^2c^2.$$

Accordingly, the equation of motion must have the form

$$W = \alpha_1 P_x + \alpha_2 P_y + \alpha_3 P_z + \alpha_4 m_0 c,$$

* h is used in this sense henceforward throughout the report.

where the coefficients α_1, α_2, α_3 and α_4 appearing in this equation are operators which must satisfy the following equations:

$$\tfrac{1}{2}(\alpha_j\alpha_k+\alpha_k\alpha_j)=\begin{matrix}0 & \text{if} & j\neq k,\\ 1 & \text{if} & j=k.\end{matrix}$$

m_0

This equation of motion refers of course to a particle of mass m_0 which is free from the action of external fields. The effect of an electromagnetic field is obtained

P_x, P_y, P_z

by replacing the momentum operators P_x, P_y, P_z, and W by

$$\left(P_x+\frac{e}{c}A_x\right),\quad \left(P_y+\frac{e}{c}A_y\right),\quad \left(P_z+\frac{e}{c}A_z\right),\quad \text{and}\quad \left(W+eU\right),$$

A_x, A_y, A_z
U

where A_x, A_y, A_z are the components of an electromagnetic vector potential, and U is the electrostatic scalar potential specifying the electromagnetic field. This device was originally justified only by a frank appeal to the correspondence principle, but one of the results which emerges from Eddington's investigations, referred to below, is a more satisfactory theory of the representation of the electromagnetic field in the quantum equation of motion.

The main results deducible from the linear wave equation constructed by Dirac relate first to the general question of the nature of the angular momentum of an electron, and secondly to the special question of the fine structure of the lines in the spectrum of atomic hydrogen. As regards the first point, it appears from Dirac's investigations that the total angular momentum of an electron consists of two parts, one of which arises from the motion of the electron as a whole in its orbit about the nucleus and is called the orbital angular momentum, while the other arises presumably from the rotation of the electron about an internal axis and is called the spin angular momentum. The magnitude of the spin angular momentum is equal to $\frac{1}{2}h$, and it differs from the orbital angular momentum in the manner in which it contributes to the total magnetic moment of the electron. The magnetic moment due to the orbital angular momentum is obtained by multiplying the latter by the factor $e/2m_0c$, as may be easily verified by considering the simple case of an electron describing a circular orbit with constant speed. Dirac's theory indicates that the magnetic moment due to the spin momentum is obtained by multiplying the latter by the factor e/m_0c. This anomaly in the ratios of the magnitude of magnetic moment to mechanical momentum for the orbital and spin components is of course in complete agreement with the semi-empirical values assumed by Uhlenbeck and Goudsmit.

Without entering into the full details of the results obtained from the fine structure on the hydrogen spectrum, it is sufficient to say that Dirac's theory reproduces Sommerfeld's formulae for the energy levels of atomic hydrogen and for the changes in these levels which are produced by an external magnetic field. The theory can also be applied in a qualitative manner to atoms with one valency electron, and it furnishes a general explanation of the fine structure of the spectral lines and of the Zeeman effect. The application of the theory to atoms with more than one valency electron is a much more difficult problem. Gaunt has attacked the problem of helium, which has two valency electrons, but the calculations are long

and difficult. Qualitative methods have been devised for the treatment of many electron atoms and the results are in general agreement with the evidence of spectroscopy.

In Dirac's theory the nucleus is assumed to be at rest—an assumption which is obviously a good approximation to the truth having regard to the ratio of the masses of a proton and electron. This assumption unfortunately ignores the fact that the nucleus probably has a spin as well as the electron. Just as the spin of the electron has been shown to be responsible for the fine structure of spectral lines, so the spin of the nucleus is thought to be the cause of the hyperfine structure. There is, however, at present no satisfactory quantitative theory of this effect. In fact the progress from the one body to the two body problem in quantum theory appears to be just as difficult as the transition from the two body to the three body problem in classical dynamics.

The principle of physical abstraction. Mathematical physics from Newton to Einstein has been dominated by the concept of a "system," i.e. a set of physical entities which can be studied without reference to the other physical entities not included in this set. The full development, however, of all the implications of this concept has only been attempted with the advent of the quantum theory. It appears that in the idea of a system the following ideas are implicitly contained at least in a general and qualitative manner—the principle of uncertainty, the theory of private and public times, the idea of absolute simultaneity, the interaction of two distinct systems, and the theory of proper mass.

The idea of a system appears to furnish the epistemological ground of the principle of uncertainty. For it is clear that a finite intelligence can only gain knowledge of the external world by isolating systems from their immediate environment, and it is equally clear that the knowledge so gained is inevitably of a partial and limited character. To know everything about something implies knowledge of everything about everything. Thus there appears to be a natural limitation on the extent of our knowledge of any physical system. This line of thought seems to converge towards the same conclusion as that reached in § 3 from the study of the nature of the actual processes of measurement in microphysics, viz. that a complete knowledge of all the characteristics of the fundamental physical systems is necessarily beyond our comprehension. It is at present premature to say whether this line of thought can be developed still further to yield the definite and quantitative form of the principle of uncertainty which is provided by Heisenberg's exchange relations.

The theory of public and private times is an immediate application of the idea of a system, and has played an important part in Eddington's theory of the interaction of two systems. Mathematical treatment of the historical development of a system is only possible when the system is treated as a whole and when the changes in the system are referred to a proper time which is a variable, independent of the co-ordinates specifying the system. (It is this fact, so fundamental and so obvious, which has completely vitiated all attempts to construct a relativistic dynamics of rigid bodies.) In the case of a composite system S which is composed of sub-systems S_1, S_2, ..., it is therefore necessary to distinguish between the public time S S_n

of the system S as a whole and the private times of its component parts S_1, S_2 and so on. By treating the system S as a whole and referring its changes to its public time we implicitly introduce simultaneity in different parts of the system and this simultaneity is of absolute significance. But on breaking up the system into its component sub-systems, we have to consider the private times of each sub-system. If these sub-systems are particles their private times are presumably subject to the relativistic transformations. Thus the concept of absolute simultaneity can be reconciled with the principle of relativity. In particular it must be noted that the universe as a whole furnishes the perfect example of a system which is both really isolated and really self-complete. Hence there is a cosmic time measuring the changes in the whole universe and providing an absolute standard of simultaneity. It is one of the paradoxes of the quantum theory that for some purposes the simplest system to consider is the whole universe.

According to Eddington[6] there is no such thing as the action of one system on another. In the crude concepts of classical physics there is a direct influence exerted by one system upon the development of another, but Eddington's theory is much more subtle. He considers a system S composed of two or more sub-systems, S_1, S_2. These sub-systems will possess their own public times, and as components of the complete system S they will fall also under the public time of S. Hence when we pass from a consideration of two isolated systems, S_1, S_2, to the single system S, composed of S_1 and S_2, the private times of S_1 and S_2 are eliminated and replaced by the public time of S. According to Eddington the transformation involved in this replacement modifies the equations of motion of S_1 and S_2 in such a way as to introduce the terms usually recognized as the interaction energy.

It is clear that the theory of a single isolated particle is one which must always be hopelessly out of touch with physical reality. As long as a system is isolated it is unobservable; on the other hand, as soon as a system emerges from isolation it loses to some extent its individuality. A physically significant theory must be the theory of some composite system. The simplest composite systems are (i) the system consisting of two particles, and referred to in the last paragraph, and (ii) the system consisting of one specified particle and the rest of the universe. By considering this second system Eddington arrives at the concept of proper mass, and furthermore is able to establish a relation between the mass of a fundamental particle and the mass of the universe as a whole. (This latter appears in his theory in a disguised form as the cosmical constant measuring the natural radius of curvature of space.) Brief reference is made to this theory in the next section. It appears to imply the existence of two kinds of fundamental particles with masses in the ratio of 1847·6 to 1, i.e. approximately the ratio of the mass of the proton to the mass of the electron.

It must be confessed that the more specific and detailed numerical deductions of Eddington's theory are much more precarious and difficult to follow than those which lead to the more general and qualitative results. However, the theory is so promising and fascinating in its unification of philosophical principles with the specific numerical values of certain physical constants, that one naturally hopes for further fruitful developments in this field.

§ 5. THE QUANTUM THEORY OF THE ULTIMATE ELEMENTS OF MATTER

The object of this, the concluding section of the report, is to give an account of the recent investigations on the properties of the ultimate elements of matter which have been inspired by quantum theory. The investigations fall into two sharply distinguished classes; first, the researches initiated by Dirac on the properties of the electron, the positron and, indirectly, the neutron; and secondly, the geometric theory constructed by Eddington which threatens to revolutionize the whole of the quantum theory. Eddington's theory discusses the nature of the interaction between electric charges, the origin of mass, and the nature of the distinction between protons and electrons, and it discusses these profound problems not only in a general and qualitative manner, but also in such a way as to yield precise quantitative estimates of such fundamental physical numbers as the fine-structure constant, the cosmical constant (which is associated with the curvature of the universe), and the ratio of the masses of the proton and electron.

The positron. Dirac's theory of the wave equation of the spin electron has been sketched in the preceding section, and some indication has been given of the remarkable success of this equation in correlating those properties of the electron which are loosely described as being due to its "spin." It will, however, be obvious from the mode of construction of Dirac's linear wave equation that it ascribes to the electron certain properties which have no analogue whatsoever in classical dynamics. These properties arise from the fact that the linear equation treats the energy as a component of a tensor of the first rank on the same footing as the three components of linear momentum, and therefore allows the energy to be either positive or negative. In fact, the solutions of the wave equation fall into two classes, according as the kinetic energy is greater than mc^2 or is less than $-mc^2$. The difficulty cannot be overcome by simply ignoring the solution which ascribes to the electron negative energy, for in the presence of an external field there will be transitions from states of positive kinetic energy to states of negative kinetic energy, and *vice versa*.

Dirac[7] has proposed an exceedingly bold and striking solution to this difficulty. According to his theory the states of negative energy have just as much physical significance as the states of positive energy; it is also admitted that transition freely takes place between these states. The crucial assumption of the theory is that in point of fact nearly all the states of negative energy in the universe are already occupied, so that the probability of a transition from a state of positive energy to a state of negative energy is enormously reduced by the smallness of the number of available empty states of negative energy. According to some expositions of the theory a uniform ocean of electrons with negative energy would be unobservable, only the small departure from exact uniformity producing an observable electromagnetic field. This statement is, however, obviously false if we ascribe to the electrons occupying the states of negative energy the same electrical properties as to the electrons occupying the states of positive energy. In Dirac's theory it is necessary to assume explicitly that this is not the case, i.e. that an electron abruptly

changes some of its electrical properties when it enters a state of negative energy. In fact, it is necessary to assume that the effective density of electrons which produces the electromagnetic field is not the actual infinite density, but the density which is due to the slight departure from the normal state of electrification of the world, in which state it is supposed that every state of negative energy is occupied and every state of positive energy is empty.

From these assumptions it is possible to deduce the existence of particles which will appear to the experimental physicist to have the same mass as an electron but to carry a positive charge of the same absolute magnitude as the negative charge carried by an electron. It is clear that, on the basis of the preceding assumptions, the physical effects of the vast hypothetical ocean of electrons with negative energy will be entirely due to the small departures from exact uniformity. A vacant state of negative energy in a region which is otherwise saturated with electrons will behave in much the same way as a single particle in a region which is otherwise empty. These vacant states will appear to have positive energy, for they will disappear when they are occupied by a particle with negative energy. Moreover, the motion of one of these empty states in an external magnetic field will be the same as that of a negative-energy electron, i.e. it will be the same as the motion of an electron with positive energy but carrying a charge $+e$. The theory thus indicates the existence of elementary particles of the same mass as the electron but carrying a charge of the opposite sign.

When Dirac first propounded this theory it was thought that these particles were protons and that subsequent revision of the theory would also enable it to account for the great difference in mass between an electron and proton. But since the theory has been proposed the physical world has been startled by the experimental discovery of positrons, i.e. of particles which behave exactly like an electron with a positive charge. It therefore now appears that Dirac's theory was actually a prediction of the existence of the positron which was only discovered experimentally two years later.

The neutron. During the last two years theoretical physics has been considerably embarrassed by the discovery of two other species of particles which have challenged the hitherto undisputed fundamental position of the electron and proton. The first kind of particle—the positron—has been satisfactorily fitted into the quantum theory, but the second species of particle—the neutron—still awaits its appropriate niche.

As is well known, the neutron is a neutral particle whose mass is equal to the mass of the proton. There are two obvious ways of constructing the theory of the neutron. It can be regarded as a fundamental structureless particle having the same irreducible character as the electron and proton, or it can be regarded as a combination of one electron and one proton, differing from an atom of hydrogen only in the extreme smallness of the effective distance between its two constituent particles. No substantial progress has been made in the first form of the theory, which appears to be an almost gratuitous assumption. The second type of theory can, however, be developed so as to account successfully for some of the main properties of the neutron.

Dirac's linear wave equation for one electron and one proton has no other physically significant solutions beyond those which describe the stationary states of ordinary atoms of hydrogen. In order to construct the theory of the neutron, therefore, it is necessary to introduce some modifications into this equation. The most natural way of doing this without introducing any *ad hoc* assumptions is to replace Dirac's first-order wave equation by the equivalent equation of the second order. When this is done it is found that the second-order wave equation possesses as solutions, not only those solutions of Dirac's first-order equation which represent atomic hydrogen, but a set of solutions whose energies can range from $-mc^2$ to $+mc^2$, including the lower limit but excluding the upper limit. (The zero of energy is here chosen so that the energy levels of H are given by the formula

$$E = mc^2 (1 - \alpha^2/2n^2),$$

where α is the fine-structure constant $= \frac{1}{137}$, and n is a positive integer.) α, n

It has been suggested by Temple[8] that the set of solutions refers to the neutron. This hypothesis implies that in its stationary state the binding energy of the neutron is approximately $2mc^2$ (this is the difference between the energy-level of the neutron in its normal state of lowest energy and the energy of H in *its* normal state). This value of the binding energy is approximately equal to one million electron volts, a value which is in good agreement with the deductions from experiment. It also appears from the theory that the serial quantum number l of the neutron is always 0, and the inner quantum number j is always $\frac{1}{2}$. This implies that the electron in the neutron has a spin of $\frac{1}{2}h$. Since the spin of the proton is also equal to $\frac{1}{2}h$, the total spin of the neutron must be either h or zero. This conclusion appears to be at variance with Heisenberg's theory of the construction of nuclei, in which it is assumed that the neutron has a spin of $\frac{1}{2}h$.

The outstanding difficulty of the theory is that it does not explain why all hydrogen atoms have not spontaneously collapsed into neutrons with the emission of radiation amounting to $2mc^2$ for each such transition. A partial explanation is that since the serial quantum number of a neutron and of a hydrogen atom in its normal state are both equal to zero, the normal state of a hydrogen atom is therefore a metastable state such that the atom cannot collapse into a neutron. This result, however, is clearly far from being a complete explanation of the rarity of the transitions from hydrogen atoms to neutrons.

Eddington's geometrical theory. The process of the geometrization of the quantum theory, as it is carried out by Eddington, starts from the simple observation that the form of the quantum theory due to Heisenberg and Dirac describes two kinds of interaction between two electrons, namely, the electromagnetic interaction due to the electric and magnetic fields to which the electrons give rise and in which they move, and also the curious form of interaction due to the operation of Pauli's exclusion principle. According to this principle, no two electrons can ever be in the same quantum state, i.e. the wave function which describes a system of electrons must be always antisymmetric in the quantum numbers which describe the different electrons. The existence of a dual interaction seems improbable. In fact it is

possible to deduce the exclusion principle from the nature of the electromagnetic interaction, for this interaction varies inversely as the distance between the electrons, and infinities in the wave equation can therefore be avoided only by adopting as solutions functions which are antisymmetric in the co-ordinates of the two electrons. Eddington's researches pursue the converse idea—that the electromagnetic interaction should be deducible from the exclusion principle[9].

The interchange of the identity of two electrons does not effect any external physical phenomenon any more than a rotation of the axes of space and time, but it differs superficially, at least, from the latter kind of transformation in appearing to us to be a discontinuous change of identity. The starting point of Eddington's investigations is the expression of the discontinuous interchange in the identity of two electrons as the final result of a series of continuous transformations. The possibility of expressing interchangeability in this way can be realized by the following crude and elementary argument. In dealing with two similar particles quantum theory has to discuss not merely the changes in the positions or momenta of these two particles, but also the probability of interchange. This probability must be associated with
θ some co-ordinate, say θ, which does not appear in the ordinary classical theories. Again, the wave function of two electrons must be written in the form

ψ
$$\psi(q, q'),$$

q, q' where q and q' summarize the co-ordinates of each electron. If the wave function of ψ is antisymmetric it will satisfy the equation

$$\psi(q, q') = -\psi(q', q).$$

This discontinuous transformation can be considered as the final result of the continuous group of transformations represented by the equation

$$\psi(q, q') \to e^{i\theta}\psi(q, q')$$

by taking θ equal to π.

Such considerations as these suggest the need for the introduction of an identity co-ordinate θ, and suggest also the way in which it will enter the theory. The identity co-ordinate θ will of course be an ignorable co-ordinate like the angular co-ordinates in the theory of one electron. Like these co-ordinates it will enter the wave equation only in the form of its conjugate momentum which will presumably be represented by a differential operator of the form

$$\frac{h}{iR'}\frac{\partial}{\partial\theta},$$

It appears that the extra term thus introduced into the wave equation is equivalent to the term which is usually introduced by an appeal to the correspondence principle and which represents the interaction energy of two particles.

The interaction of electric charges, and the fine-structure constant. The preceding account of the theory of interaction is of an extremely crude and simple character. For the detailed development of the theory the reader must be referred to Eddington's original papers[10]. However, the following considerations may be of some use in indicating the general line of argument which is there pursued. The continuous transformation with which we have replaced the discontinuous interchange of the

identity of two similar particles must lead to an interchange not only of their ordinary spatial co-ordinates but also of their associated spin variables. To allow a continuous process of interchange it is obviously necessary to supplement the physical space in which the two electrons are situated by adding extra dimensions through which the transfer can take place continuously. According to Eddington's analysis the total number of dimensions which it is necessary to introduce is 136 and the length of the path described by both particles during the process of interchange is $2\pi R$, where R is the radius which occurred above in the operator representing the identity momentum, and this distance is equal to $136 \times 2\pi r$, where r is the geometric distance between the two particles. Hence that part of the identity momentum which is due to the interchange of the spatial co-ordinates is represented by the operator

$$\frac{h}{136i} \times \frac{\partial}{r\partial\theta},$$

and this has as its fundamental proper value $h/136ir$. A more detailed analysis of the problem shows that the number 136 should probably be replaced by 137. As to the other part of the identity momentum which arose from the interchange of the spin variables, Eddington shows that this depends upon an operator of precisely the same form as that which had previously been adopted by Dirac on semi-empirical grounds, as representing the exchange energy of two similar particles. Eddington's expression for the interaction energy differs from that given by Dirac only in the fact that in Eddington's expression the fine-structure constant α which equals e^2/hc is replaced by the number $\frac{1}{137}$. (Here, e is the electronic charge, c the velocity of light *in vacuo*, and h equal to Planck's constant divided by 2π.) It is of course well known that the fine-structure constant is approximately equal to $\frac{1}{137}$. Eddington's theory attempts to indicate the theoretical reason for this value and to prove that this estimate of the value of the fine-structure constant is not merely an approximation but mathematically exact.

The origin of mass, and the cosmical constant. The preceding line of thought can be developed still further so as to deal not merely with two interacting particles but with one electron interacting with the whole of the rest of the universe[11]. The practical issue of this research is a numerical estimate of the value of the cosmical constant, i.e. of the natural radius of curvature of the universe. An elementary form of the argument is as follows: The interaction energy for one electron in the presence of another is of the form

$$I = \frac{e^2}{hc} \cdot \frac{P}{r},$$

where P is the operator referred to in the preceding paragraph and having proper values which are known to be ± 1. It follows that the interaction energy for one electron in the presence of a universe of N electrons is of the form

$$I_N = \frac{e^2}{hc} \Sigma \frac{P}{r},$$

where the summation is extended over all the N particles. It may be shown that the proper values of $(\Sigma P/r)^2$ are of the form $\pm N/\mathfrak{R}^2$, where $\mathfrak{R}$ is a species of average of $\mathfrak{R}$

r taken all over the universe, and is therefore of the same order of magnitude as the radius of curvature of the universe. It follows that the magnitude of this proper value of I_N is $\frac{e^2}{hc}\frac{\sqrt{N}}{\mathfrak{R}}$. It is suggested by Eddington that this interaction energy I_N yields that term in the wave equation which we have hitherto attributed to the proper energy mc^2 of an electron. This view leads to the equation

$$\frac{mc}{h} = I_N = \frac{e^2}{hc}\frac{\sqrt{N}}{\mathfrak{R}}.$$

By combining this result with one of the deductions from Einstein's cosmological theory it is possible to obtain an estimate of the value of the cosmical constant and of the number of particles in the universe. According to Einstein's theory we have the relation

$$GM_0/c^2 = \tfrac{1}{2}\pi\mathfrak{R},$$

G, M_0 where G is the constant of gravitation and M_0 the total mass of the universe, equal to M NM, M being the mass of a proton. Hence

$$N = \pi c^2\mathfrak{R}/2GM \quad \text{and} \quad \mathfrak{R} = \left(\frac{\pi}{2GM}\right)\left(\frac{e^2}{mc}\right)^2.$$

The numerical value of $\mathfrak{R}$ is 328 million parsecs and, as is explained in the article by Dr McVittie, this value can be used to estimate the speed of recession of the spiral nebulae. The theoretical value thus deduced for the speed of recession is 528 km. per second per megaparsec. This number is in fair agreement with the experimental results, and having regard to the precarious nature of the inference implied it is surprising that the agreement should be as close as it is.

The theory of the proton. The final and most striking achievement of Eddington's theory is the deduction of the ratio of the masses of the proton and electron[12]. The numerical identity which forms, as it were, the germ of the theory is that the ratio of the mass of the proton to that of the electron is approximately equal to $(136)^2/10$. It is not easy to summarize in a few words the nature of the arguments implied by Eddington. These arguments consist of an exceedingly bold application of the principle of physical abstraction discussed in the preceding section, and they are essentially a comparison of the sub-system which consists of an isolated electron, the sub-system consisting of macroscopic space-time, and the joint system composed of these two. The theory issues in a wave equation which is not the wave equation of a proton *and* an electron, but is the wave equation of an entity which may be either a proton *or* an electron. From this equation one can deduce a relation connecting the masses m_1 and m_2 of the proton and electron. The relation is

$$10m^2 - 136m + 1 = 0,$$

where m is the ratio of the mass of the proton or electron to a certain standard mass introduced for purposes of comparison. It follows that the ratio of the mass of the proton to the mass of the electron is equal to 1847·6, a value which is in fair agreement with experiment.

REFERENCES

(1) J. v. NEUMANN. *Z. Phys.* **57**, 30 (1929).
(2) W. HEISENBERG. *The Physical Principles of the Quantum Theory* (Cambridge University Press, 1930).
(3) P. A. M. DIRAC. *The Principles of Quantum Mechanics* (Oxford University Press, 1930).
(4) G. TEMPLE. *The General Principles of Quantum Theory* (Methuen, 1934, Chap. 2).
(5) J. v. NEUMANN. *Mathematische Grundlagen der Quantenmechanik* (Springer, 1932).
(6) A. S. EDDINGTON. *Proc. roy. Soc.* A, **126**, 696 (1930).
(7) P. A. M. DIRAC. *Proc. roy. Soc.* A, **126**, 360 (1930).
(8) G. TEMPLE. *Proc. roy. Soc.* A, **145**, 344 (1934).
(9) A. S. EDDINGTON. *Proc. roy. Soc.* A, **122**, 358 (1929).
(10) A. S. EDDINGTON. *Proc. roy. Soc.* A, **138**, 17 (1932).
(11) A. S. EDDINGTON. *Proc. roy. Soc.* A, **133**, 605 (1931).
(12) A. S. EDDINGTON. *Proc. roy. Soc.* A, **134**, 524 (1931), **143**, 327 (1934).

THE SPIRAL NEBULAE AND THE EXPANSION OF THE UNIVERSE

By G. C. McVITTIE, PH.D., University of Liverpool

§ 1. THE OBSERVATIONAL MATERIAL

THE name "spiral nebulae" is applied to certain faint telescopic objects which, in some cases, can be resolved into stars surrounding what appears to be a mass of glowing gas. The shape of the whole nebula may be spherical, elliptical or, more characteristically, of a disc-like form with arms or trailers of star-clouds winding in spirals round the central core. Nebulae of this latter type resemble our own galaxy in general shape and structure, although they fall far short of it in size. In recent years the study of the distances and spectra of these objects has attracted much attention and has led to some remarkable discoveries.

The distances. The spiral nebulae lie far beyond the regions where the methods of distance-determination used for the stars of our galaxy can be applied, one method excepted. This is the method of deducing the distance of a star from the difference between its true and its apparent brightness. As a preliminary, therefore, it is necessary to explain how the brightness of a star or of a nebula is measured and how a knowledge of the brightness leads to the evaluation of the distance.

The apparent brightness of a star can be measured visually by comparing it with an artificial star or with a set of fundamental stars. It is measured by a number called the *apparent magnitude* of the star and the scale of magnitudes is so adjusted that a first-magnitude star is $\sqrt[5]{100}$ or 2·512 times as bright as one of the second magnitude and so on, the ratio between each pair of magnitudes being the same. It turns out that on this scale really bright objects have negative magnitudes. For instance, the full moon is of apparent magnitude $-12{\cdot}5$; the planet Venus about $-4{\cdot}0$; Sirius, $-1{\cdot}58$; whilst the pole star is of about apparent magnitude $+2{\cdot}0$. Similarly *photographic apparent magnitude* can be defined by reference to the relative diameters, or the degrees of blackness, of the images of stars on the photographic plate. The photographic and the visual scales are calibrated so that stars of a certain spectral type have the same magnitudes in each. Stars of all other types then show systematic differences between photographic and visual magnitudes due to the different sensitivities of the eye and plate to light of different wave-lengths.

It is clear that the apparent magnitude of a star depends not only on its intrinsic brightness or luminosity but also on its distance. A so-called *absolute magnitude*, either photographic or visual, is therefore defined. This is the magnitude which would be assigned to the star if it were at a distance of 10 parsecs* from the sun.

m, M D, l_m We can connect the apparent magnitude m, with the absolute magnitude M, and the distance D of the star as follows. Let l_m be the total quantity of energy received

* 1 parsec = 3·258 light years = $1{\cdot}92 \times 10^{13}$ miles = $3{\cdot}09 \times 10^{13}$ km.

per second in a beam of light from the star and let l_0 be the corresponding quantity from a star of zero apparent magnitude. Also let L be the energy which would be received if the star were at a distance of 10 parsecs. Then, by the definition of apparent magnitude, we have

l_0
L

$$l_0/l_m = (2{\cdot}512)^m \qquad \text{.....(1)}$$

and, by the inverse square law for the diminution of light-intensity, we have

$$l_m/L = 10^2/D^2.$$

Also from equation (1) it follows that

$$m = 2{\cdot}5 \log_{10} (l_0/l_m)$$

and

$$M = 2{\cdot}5 \log_{10} (l_0/L).$$

Hence

$$M - m = 2{\cdot}5 \log_{10} (l_m/L) = 5 \log_{10} (10/D),$$

and finally

$$\log_{10} D = 0{\cdot}2\,(m - M) + 1 \qquad \text{.....(2).}$$

This, then, is the formula giving the distance in parsecs of a star whose apparent and absolute magnitudes are known. It is the basis of the determination of the distances of the spiral nebulae, the general principle being to find D by estimates of M and a knowledge of m.

There are three possible methods[1] for finding the distance of a nebula: (i) by identification of individual stars of known types involved in the nebula; (ii) by the apparent magnitude of the brightest stars seen in the nebula; and (iii) by the total luminosity of the nebula itself.

(i) The stars which are sought for are principally the Cepheid variable stars, of which a typical one is the star δ Cephei belonging to our own galaxy. The brightness of a star of this kind varies through a fixed range of magnitude with a fixed period, and the rise and fall in brightness is of a peculiar and easily recognized kind. It has been found[2, 10] that the Magellanic Clouds, two isolated nebulae in the southern heavens, contain a large number of these Cepheid stars with periods ranging from fifteen hours to over a hundred days. Since it is clear that these two nebulae are at a great distance from us, it follows that all these stars are approximately at an equal distance from the sun. Examination of their periods and apparent magnitudes has revealed a close correlation between these two quantities. This result leads to the conclusion that the absolute magnitude must be correlated with the period of light-fluctuation for this group of Cepheids, since the differences of apparent magnitude cannot be due to differences of distance. Assuming that all Cepheids possess this property, the coefficients of the correlation can be fixed by calculating the distance of the Cepheids of our own galaxy by some method independent of absolute-magnitude criteria. A curve in which absolute magnitude is plotted against period of light-fluctuation is thus drawn and from it the absolute magnitude of any newly discovered Cepheid, whose period is found, can be read off and its distance can be calculated by means of equation (2).

Unfortunately, Cepheid stars have only been found so far in ten nebulae[1] which are our nearest neighbours in space. They also contain novae, bright helium-line stars, and other stars, whose absolute magnitudes also are roughly known. These

stars therefore provide a rough check on the distances deduced from the Cepheids. The nearest nebula of this group, the Large Magellanic Cloud, lies at $2{\cdot}62 \times 10^4$ parsecs from the sun, and the most distant, the nebula M 81, at $7{\cdot}3 \times 10^5$ parsecs. Two important deductions are made from the observational material: (1) that the absolute visual magnitude M_n* (vis.) of an average nebula is $-14{\cdot}7$; (2) that the absolute photographic magnitude M_s (pg.) of the brightest stars of all kinds involved in the nebulae is $-6{\cdot}1$. The deviations from the mean in the case of the stars are particularly small, being $\pm 0{\cdot}04$.

M_n (vis.)
M_s (pg.)

(ii) When stars can still be identified in the nebula but the great distance makes it impossible to determine their types, resort is made to the assumption that stars have a maximum luminosity. This hypothesis is in accordance with current theories of stellar constitution and also with the empirical result found in (i). The distances of sixty-seven[1] more nebulae have been found in this way. The mean values adopted are M_n (vis.) $= -14{\cdot}9$ and M_s (pg.) $= -6{\cdot}0$ which give smaller deviations than those found in (i). The latter value inserted into equation (2) gives the formula

$$\log_{10} D = 0{\cdot}2m_s + 2{\cdot}2$$

for the distance in terms of the apparent photographic magnitude of the brightest stars involved. The most distant nebula to which this method has been applied lies at a distance of $2{\cdot}5 \times 10^8$ parsecs from the sun, but the estimated error in distances derived by this method is 20 per cent.

(iii) For still greater distances when stars can no longer be detected in the nebula, the only available method[1] is to assume that the average absolute visual magnitude of nebulae is of the order of $-14{\cdot}9$ and so deduce the distance. It is obvious that this method is reliable only when applied to a group of nebulae which are physically related. It is then fairly safe to assume that the most frequent apparent magnitude found amongst the members of the group corresponds to $M_n = -14{\cdot}9$. The orders of magnitude of the distances of twenty-five groups[9] of nebulae so established give 33×10^8 parsecs as the maximum.

Displacement of the spectral lines. With the exception of two or three of the nearest nebulae, these objects all present the phenomenon of a shift of their spectral lines towards the red end of the spectrum[3]. The displacement is very marked and is easy to observe. With the 100-inch telescope at Mount Wilson, together with a spectrograph of small dispersion (418 Å./mm. when $\lambda = 4500$), it is possible to photograph the spectra of even the most distant nebulae of class (iii). Four lines, H, K, $H\delta$ (λ 4101) and the G band (λ 4303–14) are usually clearly shown, and determinations for sixty-four nebulae are now available. The displacements in wavelength are usually listed as the corresponding Doppler velocities, and the greatest velocity so far found is 19,700 km./sec. for a cluster of nebulae in Leo. The empirical relation[1] between the logarithms of the velocities and the apparent magnitudes of the nebulae is found to be

$$\log_{10} v = 0{\cdot}2m_n + 0{\cdot}507 \pm 0{\cdot}012$$

* The subscripts n and s attached to magnitude symbols refer, respectively, to the nebula as a whole and to the brightest stars involved in the nebula.

n, s

and is closely adhered to. This formula together with formula (2) gives

$$\log_{10}(v/D) = 0{\cdot}2M_n - 0{\cdot}493 \qquad \ldots\ldots(3).$$

But $M_n = 2{\cdot}5 \log_{10} l_0 - 2{\cdot}5 \log_{10} L.$

Hence

$$\log_{10}(v/D) = (constant) - 0{\cdot}5 \log L \text{ or } v/D = constant/\sqrt{L} \qquad \ldots\ldots(4).$$

This is therefore a linear relation between distance and velocity since it seems fairly well established that the luminosity L of the nebulae is statistically constant. Introducing M_n (pg.) $= -13{\cdot}8$ into (3), we obtain

$$v = 558 \text{ km./sec. per } 10^6 \text{ parsecs} \qquad \ldots\ldots(5).$$

It is believed that the error in this value is definitely less than 20 per cent and probably not more than 10 per cent.

Distribution in space. The above results may appear to be based on very slender material when it is remembered that the Harvard survey[4] of nebulae down to the 18th apparent magnitude now in progress has already counted more than seventy-five thousand nebulae. But it is believed that the nebulae from which the result (5) is derived are very fair samples of the whole. This survey has also revealed conspicuous non-uniformities[10] in the distribution of nebulae in space as well as large numbers of close clusters of nebulae. But if sufficiently large regions of space are considered, it appears that the nebulae are on the average evenly spread out and that their linear dimensions are usually small compared with their distances apart. No really reliable estimate can yet be made of the total mass of a nebula.

§ 2. THEORETICAL INVESTIGATIONS

The theoretical problem presented by the system of the spiral nebulae may be crudely stated as follows: to construct a theory which will describe a set of particles which, though evenly spread out in space, exhibit a velocity of recession depending linearly on distance when viewed by an observer on one of them. These two properties of the set of particles are of themselves not sufficient to build up a complete theory: other assumptions with regard to the set have to be made. The chief of these are (1) that an observer on each particle must see it as the centre of the set and observe the recession phenomenon; and (2) some system of mechanics must be introduced in order to calculate the accelerations and the paths of individual particles, the density in space of the set as a whole, and so on. Ultimately, the theories proposed differ because of different assumptions under (2).

Historically the first theory[5] proposed was based on general relativity. This theory, it will be recalled, lays down as the basis of mechanics that there is a connection, expressed by Einstein's gravitational equations[6], between the material content of any region and the character of space and time appropriate to that region. Space and time themselves are inter-related and form a four-dimensional manifold which has a geometry, just as the Euclidean plane and the surface of a sphere are two-dimensional manifolds possessing geometries. The nature of the geometry is determined by the density, pressure and momentum of matter, and expressed in the

gravitational equations. Moreover, it is also assumed that the paths of free particles are certain paths called the "geodesics" of the space-time manifold, the accelerations and velocities of individual moving particles being thus fixed.

Applying these ideas to our set of particles representing the spiral nebulae, we deduce that the requirement that the particles be evenly spread out in space means that in our space-time manifold every point must be equivalent to every other. Again, since each particle sees all the rest fleeing from it, the character of the set varies with the time and so must our geometry. Manifolds with these properties are known from purely geometric theory [7]. They are characterized by possessing metrics of the form

$$ds^2 = dt^2 - \frac{e^{\beta(t)}\,(dx_1^2 + dx_2^2 + dx_3^2)}{(1 + r^2/4R^2)^2} \qquad \text{.....(6),}$$

x_1, x_2, x_3, r where $x_1^2 + x_2^2 + x_3^2 = r^2$ and t is the time, whilst x_1, x_2, x_3 are spatial coordinates.

R $1/R^2$ is a constant, positive, negative or zero, which defines what is called the curvature of space as distinct from space-time. The gravitational equations then give the

ρ, p following relations for the density ρ and pressure p of the particles occupying the region whose metric is expressed by equation (6):

$$8\pi\rho = -\lambda + \tfrac{3}{4}\left(\frac{d\beta}{dt}\right)^2 + \frac{3}{R^2}\,e^{-\beta} \qquad \text{.....(7),}$$

$$8\pi p = \lambda - \frac{d^2\beta}{dt^2} - \tfrac{3}{4}\left(\frac{d\beta}{dt}\right)^2 - \frac{1}{R^2}\,e^{-\beta} \qquad \text{.....(8),}$$

where λ is the cosmical constant, ρ and p are expressed in the customary units [6] of relativity theory and β is the expansion factor. The pressure p is thus purely hydrostatic and is due to the random motions of the particles, so that the whole system may be supposed to be at rest in the coordinate-system x_1, x_2, x_3, t. But an observer at the origin of these coordinates, using the methods of distance-determination

u employed by terrestrial astronomers, takes $e^{\frac{1}{2}\beta(t)}r$ (or u, say) to be the distance of a particle at r. This varies with the time and can be shown to give rise to a velocity of recession, if $d\beta/dt$ is positive, which is proportional to u. Since the origin may be taken at any particle, this phenomenon of recession would be observed from any particle of the set.

Observation has so far given us only the value of $d\beta/dt$, which is deduced from the observed velocity of recession of the nebulae. There remain the constants λ and R^2 in (7) and (8). A knowledge of the latter would be equivalent to a knowledge of the total extent of the system of the nebulae: if the extent were finite R^2 would certainly be positive, but if the extent were infinite R^2 might be either infinite or negative. But nothing can be said as to the actual extent of the system of the nebulae. The cosmical constant λ remains altogether arbitrary in these cosmological applications of general relativity. In short, we can say that general relativity provides us with many alternative sets of particles with somewhat different properties but all agreeing in giving the recession phenomenon and the even distribution through space.

As an alternative theory [8], we may, following E. A. Milne, abandon the relativistic hypothesis that space and time are connected with the material content of space and return to the Newtonian standpoint that space and time are assignable independently of matter. Thus Milne adopts the flat space-time of special relativity. The set of particles he considers is made up of a group of fundamental particles which all pass through a given point simultaneously, and then move away along straight lines with constant relative velocities, together with other particles which move in such a way that each of the fundamental particles appears to be the centre of the *whole* set. The added assumption that an observer attached to each particle must see, statistically, the *same* series of configurations of the whole set, determines the density of the distribution of the particles and their accelerations, etc. Hence this assumption applied to this type of set fixes its mechanics. The observers on the fundamental particles also see all the others receding from them with velocities proportional to their distances, and the non-fundamental particles tend to cluster around the fundamental ones as time goes on. Thus Milne obtains a theory of the recession phenomenon which simultaneously makes use of flat space-time and establishes a new mechanics for the system of the nebulae.

In conclusion, it may be said that it is the unique and peculiar character of the phenomenon of the recession of the nebulae which is, at bottom, responsible for the presence of such conflicting theories. Experimentally it seems hopeless to discriminate between them, their predictions in directions other than that of the recession being almost entirely incapable of test. It seems as if we must await the discovery of some new phenomenon or the development of some new theory on the lines of Sir Arthur Eddington's speculations with regard to the cosmical constant and the masses of the proton and the electron, before we can decide in favour of one theory rather than another. At present the choice is almost entirely a matter of personal taste.

REFERENCES

(1) E. Hubble and M. L. Humason. *Astrophys. J.* **74**, 43 (1931).
E. Hubble. *Proc. Acad. nat. Sci. Philad.* **15**, 168 (1929).
(2) Russell, Dugan, Stewart. *Astronomy*, **2**, 763, Ginn and Co. (1927).
(3) M. L. Humason. *Proc. Acad. nat. Sci. Philad.* **15**, 167 (1929).
H. Shapley. *Proc. Acad. nat. Sci. Philad.* **15**, 565 (1929).
M. L. Humason. *Astrophys. J.* **74**, 35 (1931).
(4) *Ann. Harv. Coll. Obs.* **88**, No. 2 (1932).
H. Shapley. *Bull. astr. Obs. Harv.* Nos. 889 and 890 (1932).
H. Shapley. *Proc. Acad. nat. Sci. Philad.* **19**, 289 (1933).
(5) G. Lemaitre. *Mon. Not. R. astr. Soc.* **91**, 483 (1931).
Sir A. S. Eddington. *The Expanding Universe*, C.U.P. (1933).
H. P. Robertson. "Relativistic Cosmology." *Rev. Mod. Phys.* **5**, 62 (1933).
(6) Sir A. S. Eddington. *Mathematical Theory of Relativity*. Ch. IV. 2nd ed. C.U.P. (1924).
(7) H. P. Robertson. *Proc. Acad. nat. Sci. Philad.* **15**, 822 (1929).
(8) E. A. Milne. *Z. Astrophys*, **6**, 1 (1933).
E. A. Milne. *Mon. Not. R. astr. Soc.* **94**, 3 (1933).
(9) H. Shapley. *Proc. Acad. nat. Sci. Philad.* **19**, 591 (1933).
(10) H. Shapley. *Mon. Not. R. astr. Soc.* **94**, 791 (1934).

CLASSICAL PHYSICS: SURFACE TENSION*

By ALLAN FERGUSON, M.A., D.Sc., F.Inst.P.

§ 1. INTRODUCTION

The subject of surface tension might itself very well run to the limits of the space assigned to this report, even though there are no fundamental changes in outlook or method to record. There are a number of minor advances in instrumental technique and in methods of computation which merit attention, and measurements of interfacial tension and of contact angles are increasing in number and in accuracy, although there is still a serious need of reliable results in this field of the subject.

§ 2. THE CAPILLARY SURFACE

More than fifty years ago, a collaboration of J. C. Adams the astronomer and Francis Bashforth ("the priest who helped the War Office in its gunnery researches," and concerning whose identity correspondents of the Sunday papers have recently shown some curiosity) resulted in the publication of a very valuable book of tables which make possible the drawing of the outline of pendent or sessile drops of fluid which are figures of revolution. The work had originated with Bashforth, but the final method of computation—an integration of the capillary equation by a continued application of Taylor's theorem—was due to Adams. Bashforth tested the results by measurements made on sessile drops of mercury. Very little notice was taken of the volume, which, after appearing for some years in the remainder catalogues, has almost disappeared from view. Some years ago Prof. Sugden recognized the value of the tables for the purpose of the reduction of observations, and quite recently Prof. Porter has made use of them in some calculations of great interest to the experimenter. The ordinary processes of analysis are competent to provide reliable corrections for the ascent or depression of liquids in what we may term "very wide" or "very narrow" tubes—systems in which h/r is small or large compared with unity. But the intermediate stage—especially in the allied problem of pendent or sessile drops—is not without interest to the experimenter, and it is precisely in this region where orthodox analysis can give little help. Here, tables such as those of Bashforth and Adams are of very great value; unfortunately there

* A report on classical general physics which covers the past year or two has nothing very exciting to put on record, and any attempt at completeness would merely result in a reproduction, in modified form, of the columns of Science Abstracts. It seemed, therefore, preferable to concentrate on one or two topics which could be treated in fairly full detail, leaving other topics to be treated similarly in succeeding reports. In the present report certain surface-tension results are discussed.

are gaps at critical points in some of these tables, and it is a question worth considering whether time might not profitably be spent on completing these tables, or on producing a set of tables which should serve the same purpose. Prof. Porter has attacked a number of problems of the type we are considering and his study of sessile drops[1] is of very special interest. We know that, when a sessile drop of radius r is so large that h/r is negligibly small compared with unity, we may write $2\beta^2=h^2$, where β is the so-called specific cohesion and may be conveniently regarded as a shorthand expression for $\gamma/g\rho$, and h is the distance between the horizontal plane through the vertex of the meridional curve, and the parallel plane of greatest horizontal section of the drop. β has the dimensions of a length; let us therefore write the equation in the non-dimensional form $\beta^2/r^2=h^2/2r^2$. The quantities determined experimentally are h and r, and, for the values of the ratio h^2/r^2 usually met with, the equation just given is not a sufficiently close approximation, and a correction must be applied in order to obtain precise values of β^2/r^2 and therefore of γ, the surface tension desired. First-order corrections, applicable to "large" drops or bubbles, have been investigated by Worthington, by Ferguson and by Rayleigh. These formulae may be written respectively as

$$\frac{2\beta^2}{r^2}=\frac{h^2}{r^2}\left(1+0{\cdot}609\frac{h}{r}\right)^{-1},$$

$$\frac{2\beta^2}{r^2}=\frac{h^2}{r^2}-0{\cdot}609\left(\frac{2\beta^2}{r^2}\right)^{\frac{3}{2}}$$

and

$$\frac{h}{\beta}=\sqrt{2}+0{\cdot}6095\frac{\beta}{r}.$$

In all these equations the curvature at the vertex (which is the origin of co-ordinates) is assumed to be negligible. It is, even now, not uncommon to make the erroneous assumption that a drop or bubble, sensibly plane at the vertex, is large enough to permit of the use of the simple formula $\beta^2/r^2=h^2/2r^2$. It is not difficult to obtain approximate expressions for this radius of curvature, and it may be shown that the correction cannot with safety be neglected for values of h^2/r^2 greater than 0·03 or 0·04 approximately.

Porter makes use of the tables of Bashforth and Adams and of Rayleigh's equation (corrected for vertex curvature) and plots a curve of r/β against r/h. There is, as we have mentioned previously, a gap to be filled, but the gap can be covered by interpolating a curve of relatively slight curvature, the curve being chosen so as to coincide, over a reasonable range on both sides of the gap, with the parts of the curve already plotted. A curve is then drawn, showing the difference Δ between β^2/r^2 and $h^2/2r^2$, the range covered being from $h^2/r^2=0$ to $h^2/r^2=0{\cdot}25$. Evidently, having determined experimentally values of h and r, we have only to find the value of Δ corresponding to the experimental values, subtract it from $h^2/2r^2$ and so arrive at the precise value of β^2/r^2.

Prof. Porter points out that for large but finite values of the radius, the ratio β^2/r^2 is less than $h^2/2r^2$, approaching this as a limiting value; for small values of the

radius, β^2/r^2 becomes very large, approaching the value infinity, while at the same time h^2/r^2 approaches the value unity. There is therefore an intermediate value for which the simple formula $\beta^2/r^2 = h^2/2r^2$ or $2\beta^2 = h^2$ holds exactly. This occurs, in fact, when $h^2/r^2 = 0{\cdot}25$, and is, obviously, a matter of considerable interest to experimenters.

For a mercury drop this condition holds with a radius of about 0·54 cm., and it is of interest to note that Quincke's figures for a sessile drop of radius 0·6 cm., when reduced in this manner, give values of the surface tension γ agreeing to 1 part in 700 with the values obtained from capillary-tube experiments. The table provided by Prof. Porter is operative over the range $h^2/r^2 = 0{\cdot}01$ to $h^2/r^2 = 0{\cdot}25$. Below values of $h^2/r^2 = 0{\cdot}04$ or thereabouts the formulae quoted earlier may be regarded as sufficiently accurate. The greatest uncertainty concerning the value of the correction Δ appertains to values of h^2/r^2 between 0·08 and 0·18.

Prof. Porter has investigated a number of similar problems by an analysis of this kind, and we may note that he has derived curves by means of which it is possible to estimate with considerable accuracy the volume of the meniscus at the surface of a liquid—a matter of some importance in various pieces of measuring apparatus.

V Simple dynamics shows that the volume V of a liquid meniscus having a contact

i angle i is given by

$$\frac{V}{2\pi\beta^3} = \frac{r}{\beta}\left(\cos i - \frac{r}{b}\right),$$

b where b is the radius of curvature at the vertex of the meridional curve. Now, since Bashforth and Adams take b as their unit of length, the data required for the computation of V may be read from their tables, which, however, extend only to values of r/β of about 3·16. This is not far enough since this extreme value corresponds, for water, to a tube about 1 cm. in radius, and Prof. Porter again makes use, to extend the range, of Rayleigh's results for wide tubes. Taking the contact angle as zero, he plots a curve between $V/\pi\beta^3$ and r/β, once again finding a gap which can be filled graphically with reasonable certainty. Thus with a given liquid and tube the value of r/β is known and the corresponding value of $V/\pi\beta^3$ may be read from the curve and V may thus be determined. "Thus for $i = 0$ and cylindrical tubes the problem may be regarded as solved for all values of r/β." The results have been extended to contact angles of 30°, 45° and 60° and to conical tubes.

Not the least interesting of Prof. Porter's investigations in this field is a simple change which produces a marked improvement in the convergency of the series formula used in the calculation of surface tensions from the results of capillary-tube experiments. The problem has been attacked many times since Poisson first produced a solution, and the latest and most accurate formula is one due to the late Lord Rayleigh, which for a liquid of zero contact angle may be written

$$2\beta^2 = r\,[h + r/3 - 0{\cdot}1288\,r^2/h + 0{\cdot}1312\,r^3/h^2].$$

This gives β^2 in terms of the height h to which a liquid rises in a tube of radius r. Take as an example a liquid for which r/h is 0·3; the value of β^2/r^2 as determined from

the tables of Bashforth and Adams is 1·8182. The equation developed by Rayleigh gives

$$\beta^2/r^2 = \tfrac{1}{6}\cdot 0\,(1 + 0{\cdot}1 - 0{\cdot}01159 + 0{\cdot}003542) = 1{\cdot}8199,$$

differing from the tabular value by about 1 in 1000. But now take as independent variable not r/h but $r/(h+r/3)$. If we write $\xi^2 = \frac{1}{2}\,(h/r + \frac{1}{3})$, Rayleigh's equation becomes ξ

$$\frac{\beta^2}{r^2} = \xi^2\left(1 - \frac{0{\cdot}03219}{\xi^2} + \frac{0{\cdot}01103}{\xi^4} - \frac{0{\cdot}0010}{\xi^8} + \ldots\right),$$

a more rapidly convergent series. Thus, taking the same value (0·3) for r/h, we see that

$$\frac{\beta^2}{r^2} = \tfrac{1}{6}\cdot 1\,(1 - 0{\cdot}00958 + 0{\cdot}00251 - 0{\cdot}0001) = 1{\cdot}8187,$$

differing by about 5 in 18,000 from the tabular value.

§ 3. SURFACE TENSION NEAR THE CRITICAL POINT

It is a remarkable fact that there is not much in the literature of surface-tension measurement which deals with determinations of surface tension made at temperatures very close to the critical point. Most experimenters have been content to extrapolate over this very important region. An attack on the problem has recently been made by Winkler and Maass who, in a paper entitled "An investigation of the surface tensions of liquids near the critical temperature[2]," have made determinations of the surface tensions of methyl ether and of propylene over a range of surface tension between 2·5 and 0·05 dynes per centimetre. They employ two experimental methods, the ring method and the capillary-rise method, and, concerning this last method, it is well here to draw attention to a delicate point discussed by Porter[3]. A capillary tube, in the sense of surface-tension experiments, is a "narrow" tube, and it is commonly assumed that if a tube is narrow at ordinary temperatures it will remain a narrow tube at temperatures in the neighbourhood of the critical point. But, as we have seen, if we write a formula such as Rayleigh's so that the left-hand side takes on the non-dimensional form β^2/r^2, the right-hand side is developed in a series of ascending powers of r/h, and the criterion of narrowness is not the smallness of the radius r, but the smallness of the ratio r/h. In the neighbourhood of the critical point h becomes very small, and the ratio r/h may take on magnitudes which make the ordinary formulae inapplicable. In fact the tube may behave as a wide one. Prof. Porter goes on to point out that, from this point of view, the values quoted and discussed are not surface-tension values but values of $h/2r$. Using Rayleigh's formula for wide tubes, he proceeds to reduce some of the experimental values, assuming that β^2/r^2 is a linear function of the temperature. The resulting curve between h/r and the temperature shows all the characteristics which are usually assigned to the surface-tension, temperature curve. "This suggests that it is quite possible that β^2 is a linear function after all."

We may note here that if surface tension varies with temperature according to the power law $\gamma = \gamma_0\,(1-m)^n$, where m is the reduced temperature, and if we assume γ_0, n, m

C, ρ_e, ρ_v, p Macleod's equation in the general form $\gamma = C(\rho_e - \rho_v)^p$, we find at once, defining β^2 in the more general way as $\beta^2 = \gamma/g(\rho_e - \rho_v)$, that

$$\beta^2 = K(1-m)^{n(p-1)/p}.$$

For most unassociated substances we may put very approximately $n = 1{\cdot}2$, $p = 4$, and we therefore obtain

$$\beta^2 = K(1-m)^{0{\cdot}9}.$$

If the power law and Macleod's law are followed, the variation of β^2 with temperature (and hence of capillary rise) is very nearly linear, and it is not surprising that the assumption of strict linearity brings us back to a curve which has the properties of the orthodox surface-tension, temperature curve. This however is incidental; the important point is the departure of r/h, at temperatures approaching the critical temperature, from the values contemplated in the equations developed for narrow tubes.

Reverting to the experiments of Winkler and Maass, we find that they tried two experimental methods—the ring method and the capillary-rise method. The last-named method is the one almost universally employed in experiments covering a wide range of temperature and, in view specially of the criticisms just advanced, it is of importance to possess results obtained by an entirely independent method. In the ring experiments, a platinum ring approximately 1 cm. in circumference was attached horizontally to the lower end of a calibrated quartz spiral of sensitivity 0·00194 gm. per mm. of extension; by means of a silk thread the spiral and ring were suspended from a glass sheath which enclosed a small iron nail. Three fine projecting glass rods were sealed to each end of the sheath in such a manner that they did not quite touch the inner walls of the glass bomb which contained this mechanism. The bomb was a cylinder of pyrex tubing, of wall-thickness 4 mm. and internal diameter 1·5 cm. It contained a shoulder on which the lower feet of the glass sheath could rest. It will be seen, therefore, that the mechanism of sheath, thread, spiral and ring could move freely along the axis of the cylindrical bomb, and the system could be raised by means of a solenoid external to the cylinder. Methyl ether was introduced by low-temperature distillation until the liquid occupied 24/85 of the bomb-space at $-78{\cdot}5°$ C. The bomb was then sealed and placed in a bath of glycoline oil. The spiral was raised until the ring was detached, as was shown by a slight jerk, the reading of the upper extremity of the spiral being taken by means of a cathetometer. A reading of the meniscus was taken, and the normal length of the spiral with the attached ring was subtracted from the difference of these readings, thus giving the extension sought. The surface tension was calculated from the relation $\gamma = mg/2l$, l being the circumference of the ring. The method was a comparative one, inasmuch as l was determined from experiments made on a liquid whose surface tension had previously been determined from a capillary-rise experiment. In the capillary-rise experiments a similar bomb was employed enclosing a capillary tube of external diameter 1 mm. and internal diameter 0·1 mm. The internal diameter of the cylindrical container was 15 mm., and it was assumed that, under these conditions, the surface tension could be calculated from the simple

l

formula $2\gamma = rhg(\rho_l - \rho_v)$. The quantitative results are not without interest, and table 1 shows a comparison of the two methods.

Table 1. Methyl ether

Temperature (° C.)	110·00	112·00	114·00	116·00	118·00	120·00
γ, ring	1·16	1·04	0·89	0·71	0·54	0·39
γ, capillary tube	1·21	1·04	0·88	0·70	0·54	0·38

These figures show the possibilities of the ring method. Nevertheless the technique was so difficult that the authors decided to rely on the capillary-rise method and the two tables following were obtained from capillary-tube observations. The authors remark that with methyl ether "the last observation which could be made with certainty was approximately 1·5° below the critical temperature," the critical temperature being presumably defined as that at which the meniscus just disappears. They point out that "the meniscus is usually defined as a *visible* line of demarcation between the vapour and the liquid phases," and that "this definition may well be extended to include dimensions not detectable by wave-lengths of ordinary light." In other words, when high accuracy is necessary, one should not define the critical temperature in terms of the properties of the observer's eye.

Table 2. Methyl ether

Temperature (° C.)	102·50	109·50	115·00	122·00	124·00	125·00
γ (dyne-cm^{-1})	1·86	1·25	0·78	0·24	0·11	0·05

Table 3. Propylene

Temperature (° C.)	62·10	73·20	77·80	83·60	88·20	89·60	90·50
γ (dyne-cm^{-1})	2·40	1·34	0·94	0·50	0·17	0·09	0·05

Of these interesting figures the authors remark that the power law for the variation of surface tension with temperature holds very accurately with n equal to 1·2. So does Macleod's equation $\gamma = C(\rho_l - \rho_v)^4$. They have calculated the free molecular surface energy E and its variation with temperature. If, following the ordinary practice we define E as $\gamma (M/\rho_l)^{\frac{2}{3}}$, and assume the Eötvös relation $E = K(\theta_c - \theta - \delta)$, the Eötvös so-called constant K varies very appreciably with temperature. If, however, we adopt the suggestion of Katayama and put E equal to $\gamma [M/(\rho_l - \rho_v)]^{\frac{2}{3}}$, it is found that the constant a in the equation

$$E = a(\theta_c - \theta)$$

is very satisfactorily constant.

The authors conclude (doubtfully) that there is a gradient in the surface-tension, temperature curve which is not zero at the critical temperature, and that a definite difference between the liquid-density and vapour-density persists above the so-called critical temperature.

In pursuance of this last remark the authors have published a most interesting experimental study entitled "Density discontinuities at the critical temperature[4]."

The matter hardly comes within the scope of this report, but it may be remembered that the well-known Cailletet-Mathias curve showing the temperature variation of the orthobaric liquid and vapour densities, quasi-parabolic in form, is usually plotted, in the two branches, to within a few degrees of the critical point, the branches being joined by a smooth extrapolated curve. It is this region which the authors have investigated, and the curve here develops a most remarkable snout or proboscis. There is not the slightest disrespect intended for the author's labours, when the present writer remarks that his first reaction to the curve was that of Mr Huckleberry Finn to the *Pilgrim's Progress*—"the statements are interesting, but steep."

§ 4. MICRO METHODS OF MEASUREMENT

Recently, interest has revived in what may be described as micro methods for the measurement of surface tension. The ordinary capillary-rise method in which the difference of level between two surfaces, one in a wide tube and one in a narrow tube, is measured, may be adapted so that a relatively small quantity of liquid is required. Richards and Coombs [5] achieve this end by placing a glass sinker in the wide tube and so effectively reducing its volume. Artifices of this type reduce the quantity of liquid needed to a few cubic centimetres, but leave untouched the problem of determining the capillary constants of a liquid of which only one or two cubic millimetres are available. One of the earliest of these micro methods is that described by Kiplinger [6]. In this method a thread of liquid is placed in a capillary tube, which is tilted until the surface of the liquid at the lower end of the capillary is plane. This position is judged by the aid of a pocket lens. Obviously, if
α, l α is the angle of tilt and l the length of the column of liquid, we have $2\gamma = rl\rho g \cos\alpha$. The results obtained are about 3 or 4 per cent in defect of the results usually accepted. The method has been modified by the present writer [7] so that the capillary tube is kept in a vertical position and pressure is applied to a short thread of liquid placed in the capillary until the lower surface of the liquid is plane. This position may be judged very accurately by watching, through a low-power microscope, the variation in the illumination of the meniscus produced by a small filament lamp. The requisite pressure is measured on a suitable manometer, and an obvious equation gives the required surface tension. The procedure is simple and the setting is quite exact, but the method, like Kiplinger's, demands a knowledge of the density of the liquid under test—not an easy quantity to determine if, *ex hypothesi*, we are restricted to one or two cubic millimetres of liquid. In the practice of the method last described, this difficulty may be overcome by taking readings of the equilibrium pressure for different lengths of liquid column and plotting a graph between pressure and length.

But a simpler way of overcoming this difficulty is to make the measurements of the equilibrium pressure required to produce a plane meniscus when the experimental tube is horizontal [8]. The deduction of the surface tension is a very simple matter if the bore of the tube is small enough to permit of the neglect of the gravita-

tional distortion of the meniscus. This was tested by means of experiments made with the tube first vertical and then horizontal, and it was found that with tubes less than 1 mm. in bore it was quite safe to equate the equilibrium pressure, as measured on a separate manometer, to $2\gamma/r$, r being the radius of the horizontal capillary. An experiment with a tube 0·046 cm. in diameter gave $73{\cdot}5_5$ for the surface tension of water at 15° C., tubes of radii 0·036 and 0·063 cm. yielding values of 73·47 and $73{\cdot}3_8$ dyne/cm. respectively at the same temperature.

Mouquin and Natelson[9] use as horizontal capillary a thermometer tube broken at both ends. A mark is etched on the upper (reservoir) bulb, a thread of liquid is drawn into the thermometer tube, and the system is connected to a manometer of the inclined-tube type and to an apparatus for regulating the pressure. The pressure is then adjusted until, one end of the thread being in the thermometer tube, the meniscus at the other end coincides with the mark etched on the bulb. As the ends of the thread are situated at points on the tube having different radii, a definite pressure-difference is required for equilibrium, which pressure-difference is easily evaluated in terms of the radii of the tube at the points in question. But it is not easy to determine accurately the bore of the tube at the mark etched on the bulb, and the authors therefore make the method a comparative one, plotting a calibration curve between manometer readings and surface tension for a number of liquids of known surface tension.

The same authors[10] describe a method which has some relation to Kiplinger's method, inasmuch as an equilibrium position is found for a thread of liquid in an open capillary. In the practice of their method, however, the capillary is held in a vertical position, and equilibrium is ensured by using a conical tube. Carlton Sutton[11] points out that, if two capillary tubes of different radii be fused together and held in a vertical position, an equilibrium position may be found for a column of liquid in the tube, by attaching one end of the tube to a pressure apparatus, and applying pressure to bring the meniscus to a definite point in the tube of finer bore; the tube is now inverted so that the finer tube is lowermost, and the position of (unstable) equilibrium of the column is found. The author finds that consistent and reproducible results which are in good agreement with the values given by independent methods may be obtained readily and rapidly.

An obvious micro method which has some advantages, especially for investigators of temperature variation of surface tension, consists in making a U tube of two capillaries of different bore, holding the tube in a vertical position, and measuring, on a suitable manometer, the pressure required to bring the two menisci to the same level in the U tube. Assume that the wider tube has twice the bore of the narrower tube and that the liquid in the manometer has the same density as that of the liquid under test. Then it is easy to see that the pressure head registered by the manometer is equal to the ordinary capillary rise of the liquid in the tube of wider bore.

The method is simple, may be made very sensitive by the choice of a suitable manometer and involves no knowledge of the density of the liquid employed. It promises to be very suitable for investigating the temperature variation of surface tension over a wide range.

§ 5. ANGLES OF CONTACT

1, 2 A_{12} If a solid be denoted by the subscript 1 and a liquid by the subscript 2, the quantity A_{12} defined by the relation

$$A_{12}=\gamma_1-\gamma_{12}=\gamma_2\cos\theta_{12}$$

is called the *adhesion tension*. The introduction of such a concept naturally makes the determination of contact angles a matter of special importance, and interest in this problem has recently revived. Bartell and Merrill[12] make direct determinations of this quantity by obtaining the photomicrograph of the liquid interface or meniscus upon a photographic plate which could be used as a lantern slide. With this, negative projections were then made on a sheet of white paper. "The contact angles were estimated by the actual drawing of tangents. The capillaries used were drawn from quartz tubing and ranged from 0·2 to 0·6 mm. in diameter. It is interesting to note that at first attempts were made to clean the tubes after drawing them... it was found that the washing of the inside walls of the tubes with acids produced undesirable changes on the surfaces and hence the practice thereafter was to use tubes just as taken from the fire in the drawing process. Glass capillaries were prepared for use in the same manner."

Schultze[13] determines contact angles from measurements of the capillary rise on single glass plates and rods from 0·4 to 2·0 mm. thick, inclined and normal to surfaces of water and mineral oils.

§ 6. MOLECULAR SURFACE ENERGY

A number of workers are now exhibiting results in terms of the total molecular surface energy as defined by Katayama. It may be remembered that Katayama altered the ordinary definition of free and total molecular surface energy by substituting the difference between the liquid and vapour orthobaric densities for the liquid density. It is commonly assumed that the march with temperature of the free molecular surface energy as ordinarily defined is linear, and that the total molecular surface energy as commonly defined is independent of the temperature. These views are far from being correct and the variation with temperature of these two quantities may be expressed by formulae which, while not giving support to the views commonly accepted, are in very close agreement with experimental values. The matter turns upon the formulae which may be deduced to represent the variation with temperature of surface tension and of density.

For unassociated liquids the power law $\gamma=\gamma_0(1-m)^n$ represents the experimental facts with considerable accuracy, where m stands for reduced temperature. If high accuracy is desired it must not be assumed that n is equal to 1·2, but the variation of n from liquid to liquid must be taken into account. If we assume this power law and also assume a generalized form of McLeod's law—that is

$$\gamma=C(\rho_L-\rho_V)^p,$$

where p may or may not be equal to 4, then elimination of γ leads to a relation

between the temperature and $(\rho_L - \rho_V)$. Combining this with the law of rectilinear diameters which states that $(\rho_L + \rho_V)$ is a linear function of m, we easily arrive at an equation which represents the variation with temperature of the orthobaric values of the liquid or the vapour density. In its simplest form which assumes that $n = 1{\cdot}2$ and $p = 4$, the resulting equation[14] is $\rho = 2\rho_c [A(1-m)^{0{\cdot}3} + (1 - 0{\cdot}5m)]$, where A is a constant approximately equal to unity. This formula represents the temperature variation of liquid density over the whole range of existence of the liquid with considerable accuracy, and affords a very close representation of the experimental facts if the values of A, of n and of p for the individual liquids are taken into account. If this expression for the liquid density be substituted in the expression for the free and total molecular surface energies, we then obtain formulae which give the temperature variation of these quantities very accurately*. The variation of the free molecular surface energy with temperature is not linear, neither is the total molecular surface energy independent of the temperature. On the contrary, this latter quantity rises to a slight but quite definite maximum, which occurs at a temperature about 0·8 of the critical temperature, thereafter falling rapidly to zero at the critical temperature. The experimental figures usually quoted in support of constancy clearly show this maximum value.

A

Katayama's assumption that the free molecular surface energy defined as $\gamma \{M/(\rho_L - \rho_V)\}^{\frac{2}{3}}$ is a linear function of the reduced temperature leads to the result that the total molecular surface energy as similarly defined increases linearly with the reduced temperature. These results are now very commonly used to express in brief manner the experimental figures obtained[15].

An equation such as $E = E_0(1 + 0{\cdot}2m)$, however, where E stands for total molecular surface energy as defined by Katayama, can hardly be expected to be valid over the whole range of existence of the liquid, leading as it does to finite values of E at the critical temperature. In fact the equation depends on the assumption that the n of the power law and the p of McLeod's law are universal constants equal respectively to 1·2 and to 4. This is not the case, although there is a significant grouping of the experimental figures round these values for many unassociated liquids. If we take into account the individual values of these constants, some results as yet unpublished show that we can represent the temperature variation of the total molecular surface energy as defined in the Katayama manner by

$$E = E_0(1-m)^x [1 + (n-1)m],$$

where x is $(n - 1 - 2n/3p)$ and is a small but usually quite appreciable positive quantity. At low values of the reduced temperature, the power term in this expression has little effect and E increases with m very approximately in a linear fashion. As

x

* If the free surface energy (surface tension) be denoted by γ, and the total surface energy by λ, then, as is well known, $\lambda = \gamma - \theta(\partial\gamma/\partial\theta)$. Hence, if the power law $\gamma = \gamma_0(1-m)^n$ represents accurately the variation of surface tension with temperature, substitution for γ and for $\partial\gamma/\partial m$ (or $\partial\gamma/\partial\theta$) in the equation for λ leads to a simple expression for the variation of λ with temperature which does not exhibit a maximum value, as it is sometimes supposed to do. If we insert these values of γ and of λ in the expressions for the free molecular surface energy $[\gamma(M/\rho)^{\frac{2}{3}}]$ and the total molecular surface energy $[\lambda(M/\rho)^{\frac{2}{3}}]$ respectively, and employ the equation which shows the march of liquid density with temperature, we easily arrive at the expressions discussed in the text.

m approaches unity, however, the curve between E and m bends over sharply and descends rapidly to a zero value at the critical temperature. Similarly if we take into account deviations of n and of p from the commonly assumed values of 1·2 and of 4, the Katayama equation for the variation of free molecular surface energy with temperature is not strictly linear; this deviation from linearity can also be expressed by means of a simple equation which is more in accord with experimental facts. The constant E_0, which may be called the zero total molecular surface energy of the supercooled liquid, is important and its properties are under investigation.

It may be noted in passing that attempts to test the McLeod law in its generalized form $\gamma = C(\rho_L - \rho_V)^p$ by making a logarithmic plot between γ and $(\rho_L - \rho_V)$ show that, for many liquids for which the power law is closely followed, there are appreciable deviations from the McLeod law in the neighbourhood of the critical temperature. Moreover, over the region in which the logarithmic plot is linear there are appreciable deviations from the value $p = 4$. These deviations are sufficient to introduce appreciable errors into the numerical value of the parachor if the common practice is followed of taking the fourth root instead of the pth root in the calculation for the parachor. Nevertheless, it should be noticed that Eucken[16] has shown that the parachor relation in the form

$$\gamma V^4 = \text{constant},$$

V where V is the molecular volume, can be derived from a general equation of state, which is applicable to liquids at low and high pressures and is in agreement with the London attraction law. The additivity of the parachor is explained by the fact that it represents about 2·9 times the zero volume.

§ 7. VARIATION OF SURFACE TENSION WITH TIME

That surface tension varies with time is a fact that has long been established, and the difference between dynamic values of surface tensions as determined, for example, by the oscillating-jet method and the static values determined by capillary-rise experiments is well known. But it was not until comparatively recently that the ageing of surfaces with time was studied in any quantitative manner. For example, Lecomte du Nouy[17], using the ring-torsion-balance method, investigated the variation of surface tension with time for certain colloidal solutions, and concluded that for a number of them the surface tension was an exponential function of the time, being proportional to exp. $[-k\sqrt{t}]$; and Johlin[18], using the capillary-rise method, found that the surface tension was proportional to t^{-n}. Bigelow and Washburn[19] obtained similar results but noted the existence of a period of inactivity. Ghosh and Nath[20] have also investigated this problem for various solutions. They follow Freundlich's advice that "the method of adhesion rings is to be particularly recommended for determining the surface tension of any plane liquid surface which it is desired to disturb as little as possible," and they therefore use a nickel ring which is attached to one arm of the beam of an ordinary balance, the other arm carrying a brass pan and a gold chain so that fine adjustments may be made in the

fashion of a chainomatic balance. The beaker containing the solution was kept at a temperature 5° below that of the room by circulating cold water through a glass trough in which the beaker was immersed. The method was used as a comparative one, the apparatus being calibrated by measurements made on pure water. The authors divide the solutions examined into three types. Type A comprises "true solutions of simple molecules whose velocity of diffusion is fairly large and the equilibrium concentration at the surface is reached too quickly to be detected by this method of measuring surface tension." Type B comprises "solutions consisting of multi-molecular aggregates having small velocities of diffusion." For this type the authors attempt a simple kinetic theory which leads to a logarithmic formula between surface tension and time. Type C comprises colloidal solutions for which the authors find "that the variation of surface tension with time is not regular—maximum and minimum (*sic*) being observed." The authors remark that this classification receives a qualitative confirmation from the results of dialysing experiments.

Nevertheless, despite the authority of Freundlich, it may be questioned whether the ring method is best suited for experiments of this type. A method devised by the present writer[21] seems more suitable. In the practice of this method a vertical capillary tube dips into the solution under test and the pressure required to force the meniscus to the bottom of the tube and to hold it there in equilibrium is measured on a separate manometer. By sending a few bubbles through the liquid a fresh meniscus may be obtained. It can be held in position without the slightest disturbance, and the ageing of the surface can be studied by observing the pressure required to hold the meniscus in its equilibrium position. Using this method one of the writer's students, Dr L. E. Sharp, has carried out a very comprehensive series of researches on the variation of surface tension with concentration and with temperature for aqueous solutions of certain fatty acids (stearic, palmitic, myristic and lauric acids) and aqueous solutions of sodium oleate. Taking first the results obtained by varying temperature at constant concentration, it is found that in every instance there is a period of inactivity over which the surface tension remains constant. At the end of this period the surface tension falls rapidly to a lower stable value, the drop being of the order of 30 c.g.s. units. At low temperatures, this period of inactivity is very great, but it decreases rapidly and in a perfectly regular manner with increase of temperature. Similar results hold for change of concentration at constant temperature, the period of inactivity decreasing with rise in concentration*.

* These results, which are as yet unpublished, form part of a thesis presented by Dr Sharp to the University of London for the degree of Ph.D.

REFERENCES

(1) Porter. *Phil. Mag.* **15**, 163 (1933).
(2) Winkler and Maass. *Canad. J. Res.* **9**, 65 (1933).
(3) Porter. *Nature*, Lond., **130**, 929 (1932).
(4) Winkler and Maass. *Canad. J. Res.* **9**, 613 (1933).
(5) Richards and Coombs. *J. Amer. chem. Soc.* **37**, 1656 (1915).
(6) Kiplinger. *J. Amer. chem. Soc.* **42**, 472 (1920).
(7) Ferguson. *Proc. phys. Soc.* **36**, 37 (1923).
(8) Ferguson and Kennedy. *Proc. phys. Soc.* **44**, 511 (1932).
(9) Mouquin and Natelson. *J. phys. Chem.* **35**, 1931 (1931).
(10) —— *Mikrochemie*, **12**, 293 (1933).
(11) Carlton Sutton. *Proc. phys. Soc.* **45**, 88 (1933).
(12) Bartell and Merrill. *J. phys. Chem.* **36**, 1179 (1932).
(13) Schultze. *Kolloidzschr.* **62**, 294 (1933).
(14) Ferguson and Miller. *Proc. phys. Soc.* **46**, 140 (1934).
(15) Morino. *Sci. Pap. Inst. phys. chem. Res.*, Tokyo, **11**, 1018–43 (1932).
(16) Eucken. *Nachr. Ges. Wiss. Göttingen*, p. 340 (1933).
(17) Lecomte du Nouy. *Phil. Mag.* **48**, 264 (1924).
(18) Johlin. *J. phys. Chem.* **29**, 271, 897, 1129 (1925).
(19) Bigelow and Washburn. *J. phys. Chem.* **32**, 281 (1928).
(20) Ghosh and Nath. *J. phys. Chem.* **36**, 1916 (1932).
(21) Ferguson and Hakes. *Proc. phys. Soc.* **41**, 214 (1929).

BURGERS'S THEORY OF TURBULENCE

By G. F. P. TRUBRIDGE

ABSTRACT. This report gives an account of Burgers's attempt to apply the methods of statistical mechanics to the theory of the completely turbulent motion of a fluid between two parallel walls. The shearing-force condition and the dissipation condition which govern the flow in three dimensions are developed by taking mean values over a distance great compared with the width of the channel. Details are then given of Burgers's method. The stream function for the relative motion in two dimensions is used to describe the probabilities of the various states of motion, the amplitudes of the modes of motion being introduced as coordinates in a phase space. An expression for the distribution function is obtained, the assumption being made that only quantities defined for instantaneously observed modes of motion are to be considered. Since the motion must be stationary over a long period of time a special condition is introduced which must be satisfied by the sequence of quantities so defined. The stream function is developed according to the proper functions of a certain differential equation of the fourth order. An attempt is made to calculate the statistical mean values of the squares of the amplitudes of the modes of motion by introducing a plausible value for an arbitrary function which occurs. The mean motion is then treated in a similar way. The results lead to values for the distribution of the mean velocity which are rather higher than the experimental values. A lower limit is found for the value of the resistance coefficient. It is shown that an important integral in the theory becomes divergent, and the suggestion is made that a deeper consideration of the significance of Liouville's theorem in the theory may lead to a restriction on the possible modes of motion, and so make the integral convergent.

§ 1. INTRODUCTION

Several theories have been put forward to explain the phenomena of turbulent motion, but not one of them is completely successful. A qualitative explanation of the increase of resistance which accompanies the appearance of the turbulent state of flow in a channel has been given by Reynolds[1] and Lorentz[2]. Prandtl[3], Von Karman[4] and Taylor[5] have recently developed theories yielding many interesting results which, it appears, are in good agreement with the last experimental results[6]. Not one of these theories however shows the structure of the turbulent motion, the connection between the motion in the central part of the channel and that in the boundary layers in the immediate vicinity of the walls; nor do the theories give any information about the balance of energy between the principal and relative motions. The suggestion has been made by several writers[7,8,9] that an adequate theory for completely turbulent flow will only be developed if the problem is treated statistically. During 1929 to 1933 Burgers[10,11] succeeded in applying the methods of statistical mechanics to the problem of the turbulent flow

of an incompressible viscous fluid between two parallel walls. Before entering into a discussion of these methods we shall first establish the two fundamental conditions which will be used to formulate the solution. These conditions are the connection between the total shearing force and the turbulent shearing force at a point in the fluid, and the dissipation condition, which determines the equilibrium of the energy dissipated by the viscosity with the energy transmitted to the relative motions from the mean motion.

§ 2. THE FUNDAMENTAL CONDITIONS

h

We consider the motion between two fixed parallel walls at a distance h apart. Provisionally we introduce three-dimensional coordinates and take the y axis perpendicular to the walls, the x axis in the direction of the main flow and the z axis perpendicular to both of them. We may imagine that the fluid is bounded by two other pairs of walls at distances great compared with h from the xy plane, and then discuss the motion in the vicinity of this plane. The motion is assumed to be completely turbulent, so that the average properties of the flow are fairly regular. The fundamental equations of hydrodynamics are taken to be valid for this type of motion. In order to simplify the equations non-dimensional variables are introduced by dividing all lengths by h and all velocities by V_0, the principal velocity of flow, that is, the volume of liquid which in unit time flows through a perpendicular section of the channel divided by the area of that section; then the pressure-gradients must be divided by $\rho V_0^2/h$, while the viscosity coefficient must be replaced by the inverse Reynolds's number $R^{-1}=\mu/\rho V_0 h$. By making use of the equation of continuity the equations of motion may be written as

V_0

ρ

R, μ

x, y, z, t

u, v, w, p

$$\left.\begin{aligned}
&\frac{\partial u}{\partial t}+\frac{\partial (u^2)}{\partial x}+\frac{\partial (uv)}{\partial y}+\frac{\partial (uw)}{\partial z}=-\frac{\partial p}{\partial x}+R^{-1}\nabla^2 u\\
&\frac{\partial v}{\partial t}+\frac{\partial (uv)}{\partial x}+\frac{\partial (v^2)}{\partial y}+\frac{\partial (vw)}{\partial z}=-\frac{\partial p}{\partial y}+R^{-1}\nabla^2 v\\
&\frac{\partial w}{\partial t}+\frac{\partial (uw)}{\partial x}+\frac{\partial (vw)}{\partial y}+\frac{\partial (w^2)}{\partial z}=-\frac{\partial p}{\partial z}+R^{-1}\nabla^2 w
\end{aligned}\right\} \quad \ldots\ldots(1),$$

∇^2, ∂

where ∇^2 denotes the operator $\dfrac{\partial^2}{\partial x^2}+\dfrac{\partial^2}{\partial y^2}+\dfrac{\partial^2}{\partial z^2}$.

The components of the velocity u, v, w at a point in the fluid are now regarded as fluctuating functions ([6], p. 364). A mean value with respect to x for a function

ϕ

ϕ of the variables x, y, z, t is introduced by means of the equation

$\bar{\phi}, L, x_0$

$$\bar{\phi}=\frac{1}{L}\int_{x_0}^{x_0+L}\phi\,(x, y, z, t)\,dx \qquad \ldots\ldots(2),$$

where x_0 and x_0+L are the end values of x, L being the length of the channel considered. Introducing mean values we obtain the relation

$$\frac{\overline{\partial u}}{\partial t}+\frac{1}{L}\left|u^2\right|+\frac{\overline{\partial (uv)}}{\partial y}+\frac{\overline{\partial (uw)}}{\partial z}=-\frac{\overline{\partial p}}{\partial x}+R^{-1}\left(\frac{1}{L}\left|\frac{\partial u}{\partial x}\right|+\frac{\overline{\partial^2 u}}{\partial y^2}+\frac{\overline{\partial^2 u}}{\partial z^2}\right) \quad \ldots(3),$$

with similar equations for the other components. The parallel-line symbol denotes the difference in the values of the quantity for $x=x_0$ and x_0+L. The quantities between parallel lines are now assumed to be small compared with the length L and are neglected. This condition necessitates that the length of the channel shall be great compared with the depth.

Following Burgers we decompose the real flow into the so-called mean flow and the continually fluctuating relative motion. We write

$$u=\bar{u}+u', \quad v=\bar{v}+v', \quad w=\bar{w}+w' \qquad \text{.....(4),}$$

$\bar{u}, \bar{v}, \bar{w}$

where $\bar{u}, \bar{v}, \bar{w}$ are the components of the mean motion and u', v', w' are the component velocities of the relative motion. This relative motion constitutes a superimposed turbulence. Certain rules which the mean quantities obey approximately when the fluctuations are sufficiently numerous and distributed at random have been developed by several investigators. These rules have been given by Bateman ([6], p. 366) in the form

u', v', w'

$$\bar{\bar{a}}=\bar{a}, \quad \overline{a\varpi}=\bar{a}\,.\,\bar{\varpi}, \quad \overline{a+\varpi}=\bar{a}+\bar{\varpi}, \quad \frac{\overline{\partial a}}{\partial s}=\frac{\partial \bar{a}}{\partial s} \qquad \text{.....(5),}$$

a, ϖ, s

where s denotes any one of the four variables x, y, z, t and a and ϖ are functions of the velocities. The last relation is true when $s=y$, z or t, provided the function a satisfies the conditions for differentiation under the sign of integration. When $s=x$, $\frac{\partial \bar{a}}{\partial x}=0$ since $\bar{a}$ is independent of x. But from (2),

$$\frac{\overline{\partial a}}{\partial x}=\frac{1}{L}\int_{x_0}^{x_0+L}\frac{\partial a}{\partial x}\,dx=\frac{1}{L}\,|\,a\,|\,.$$

Hence $\frac{\overline{\partial a}}{\partial x}=\frac{\partial \bar{a}}{\partial x}$ provided $|\,a\,|$ is not of the order L. Investigations regarding the validity of these rules for smoothing out the quantities involved in the above equations show that the relations (5) may be used with a sufficient degree of accuracy, but it is desirable that detailed consideration should be given to the assumption that these relations are valid [12, 13, 14, 15, 16]. Defining the mean velocity U by the equation

U

$$U=\bar{u}=\frac{1}{L}\int_{x_0}^{x_0+L}u\,dx, \text{ and setting } \bar{v}=\bar{w}=0 \qquad \text{.....(6),}$$

we obtain the relations

$$\left.\begin{aligned} &\overline{u'}=\overline{v'}=\overline{w'}=0,\\ &\overline{uu}=\bar{u}\bar{u}+\overline{u'u'}=U^2+\overline{u'^2}\\ &\overline{uv}=\bar{u}\bar{v}+\overline{u'v'}=\overline{u'v'}, \text{ etc.} \end{aligned}\right\} \qquad \text{.....(7).}$$

Substituting these results in the equations (3) and using the relations (5), (6) and (7) we obtain the equations

$$\left.\begin{aligned} &\frac{\partial U}{\partial t}+\frac{\partial\,(\overline{u'v'})}{\partial y}+\frac{\partial\,(\overline{u'w'})}{\partial z}=-\frac{\partial \bar{p}}{\partial x}+R^{-1}\left(\frac{\partial^2 U}{\partial y^2}+\frac{\partial^2 U}{\partial z^2}\right)\\ &\frac{\partial\,(\overline{v'^2})}{\partial y}+\frac{\partial\,(\overline{v'w'})}{\partial z}=-\frac{\partial \bar{p}}{\partial y}\\ &\frac{\partial\,(\overline{v'w'})}{\partial y}+\frac{\partial\,(\overline{w'^2})}{\partial z}=-\frac{\partial \bar{p}}{\partial z} \end{aligned}\right\} \qquad \text{.....(8).}$$

These are the conditions which govern the three-dimensional problems, but unfortunately it appears that the statistical methods can be applied only when the motion is restricted to two dimensions. According to recent views it is likely that vortices having their axes parallel to the direction of the main flow play the most important part in determining the nature of turbulent motion. A two-dimensional consideration precludes the possibility of the existence of such vortices in the fluid. Burgers* has recently suggested possible methods for attacking the problem in three dimensions.

A further condition which is used in the formulation of the solution to the problem is the condition for the dissipation of energy of the system. Reverting to equations (1) we set

P, p'
$$u = U + u', \quad v = v', \quad w = w', \quad p = P + p' \qquad \text{.....(9)},$$

where P is the mean pressure, and obtain the relations

$$\left.\begin{aligned}
&\frac{\partial u'}{\partial t} + 2U\frac{\partial u'}{\partial x} + U\frac{\partial v'}{\partial y} + U\frac{\partial w'}{\partial z} + v'\frac{\partial U}{\partial y} + \frac{\partial}{\partial x}(u'^2) + \frac{\partial}{\partial y}(u'v') + \frac{\partial}{\partial z}(u'w') = -\frac{\partial p'}{\partial x} + R^{-1}\nabla^2 u' \\
&\frac{\partial v'}{\partial t} + U\frac{\partial v'}{\partial x} + \frac{\partial}{\partial x}(u'v') + \frac{\partial}{\partial y}(v'^2) + \frac{\partial}{\partial z}(v'w') = -\frac{\partial p'}{\partial y} + R^{-1}\nabla^2 v' \\
&\frac{\partial w'}{\partial t} + U\frac{\partial w'}{\partial x} + \frac{\partial}{\partial x}(u'w') + \frac{\partial}{\partial y}(v'w') + \frac{\partial}{\partial z}(w'^2) = -\frac{\partial p'}{\partial z} + R^{-1}\nabla^2 w'
\end{aligned}\right\} \quad (10),$$

and
$$\frac{\partial U}{\partial t} + \frac{\partial}{\partial x}(U^2) = -\frac{\partial P}{\partial x} + R^{-1}\nabla^2 U, \quad \frac{\partial P}{\partial y} = 0, \quad \frac{\partial P}{\partial z} = 0.$$

The first three results are now multiplied by u', v', w' and added. The components
$\xi', \eta', \zeta', \delta E'$
ξ', η', ζ', of the vorticity of the relative motion, and the symbol $\delta E'$ for the element of the energy of the relative motion are now introduced. Using a result from vector analysis, taking mean values and using the rules of combination which the mean values obey we finally get the result

$$\begin{aligned}
&\frac{\partial \delta E'}{\partial t} + \overline{u'v'}\frac{\partial U}{\partial y} + \overline{u'\frac{\partial}{\partial y}(u'v')} + \overline{u'\frac{\partial}{\partial z}(u'w')} + \overline{v'\frac{\partial}{\partial y}(v'^2)} + \overline{v'\frac{\partial}{\partial z}(v'w')} \\
&+ \overline{w'\frac{\partial}{\partial y}(v'w')} + \overline{w'\frac{\partial}{\partial z}(w'^2)} = -R^{-1}(\overline{\xi'^2} + \overline{\eta'^2} + \overline{\zeta'^2}) + R^{-1}\overline{\frac{\partial}{\partial y}(w'\xi' - u'\zeta')} \\
&+ R^{-1}\overline{\frac{\partial}{\partial z}(u'\eta' - v'\xi')} - \overline{u'\frac{\partial p'}{\partial x}} - \overline{v'\frac{\partial p'}{\partial y}} - \overline{w'\frac{\partial p'}{\partial z}} \qquad \text{.....(11)},
\end{aligned}$$

where all the terms which are small in comparison with L are neglected.

Integrating this result in two dimensions over the cross-section of the channel, and using the facts that $\overline{p'} = 0$ and $u' = v' = 0$ at the walls, we obtain the relation

$$\frac{\partial E'}{\partial t} = -\int_{-\frac{1}{2}}^{\frac{1}{2}} \overline{u'v'}\frac{\partial U}{\partial y}\,dy - R^{-1}\int_{-\frac{1}{2}}^{\frac{1}{2}} \overline{\zeta'^2}\,dy.$$

* Private communication, June 1934.

This condition was given by Burgers[17] in a previous paper on the subject. The condition may be written in the alternative form

$$\frac{\partial E'}{\partial t} = \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{\partial}{\partial y}(\overline{u'v'})\, U dy - R^{-1} \int_{-\frac{1}{2}}^{\frac{1}{2}} \overline{\zeta'^2}\, dy \qquad \text{.....(12),}$$

since $U = 0$ at the walls.

Using the fact that

$$\frac{\partial}{\partial y}(\overline{v'^2}) = 2\overline{v'}\,\frac{\overline{\partial v'}}{\partial y} = 0 \qquad \text{.....(13)}$$

in equations (8) and taking the two-dimensional form, we see that $-\frac{\overline{\partial p}}{\partial x}$ is independent of x, y and z. For simplicity it is usual to write $2C = -\frac{\overline{\partial p}}{\partial x}$, and (8) then gives C

$$\frac{\partial U}{\partial t} = 2C - \frac{\partial}{\partial y}(\overline{u'v'}) + R^{-1}\frac{\partial^2 U}{\partial y^2} \qquad \text{.....(14).}$$

To find the rate of increase of the kinetic energy of the flow per unit length of the channel we must also obtain the energy of the mean motion. Multiplying (14) by U and integrating we obtain

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} U\frac{\partial U}{\partial t}\, dy = 2C\int_{-\frac{1}{2}}^{\frac{1}{2}} U dy - \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{\partial}{\partial y}(\overline{u'v'})\, U dy + R^{-1}\int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{\partial^2 U}{\partial y^2}\, U dy.$$

Now $\int_{-\frac{1}{2}}^{\frac{1}{2}} U dy = 1$, and $U = 0$ when $y = \pm\frac{1}{2}$; hence if E_0 denote the energy of the E_0 mean motion we have

$$\frac{\partial E_0}{\partial t} = 2C - \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{\partial}{\partial y}(\overline{u'v'})\, U dy - R^{-1}\int_{-\frac{1}{2}}^{\frac{1}{2}} \left(\frac{\partial U}{\partial y}\right)^2 dy \qquad \text{.....(15)}$$

on integrating the last integral by parts. The expression for the rate of increase of the total kinetic energy E of the flow per unit length of the channel is thus given by

$$\frac{\partial E}{\partial t} = \frac{\partial E_0}{\partial t} + \frac{\partial E'}{\partial t} = 2C - R^{-1}\int_{-\frac{1}{2}}^{\frac{1}{2}} dy \left(\zeta'^2 + \left(\frac{\partial U}{\partial y}\right)^2\right) \qquad \text{.....(16).} \quad E$$

This is the required condition for the dissipation of energy. The conditions (14) and (16) are fundamental in the theory of turbulence as developed by Burgers.

§ 3. DEVELOPMENT OF THE STATISTICAL METHOD

The application of the methods of statistical mechanics requires the introduction of some process which allows us to describe the probabilities of the various states of motion. In the case of three-dimensional motion there does not yet appear to be an appropriate quantity available, but in the case of two-dimensional motion the stream function can be used to describe the states of the motion. A stream function ψ is therefore introduced for the relative motion; with it u', v', ζ' are connected by ψ the formulae:

$$u' = \frac{\partial \psi}{\partial y}, \quad v' = -\frac{\partial \psi}{\partial x}, \quad \zeta' = \frac{\partial v'}{\partial x} - \frac{\partial u'}{\partial y} = -\nabla^2\psi \qquad \text{.....(17).}$$

As ψ must satisfy the boundary conditions: $\psi = 0$, $\partial\psi/\partial y = 0$ for $y = \pm\frac{1}{2}$, the stream function which will be developed will be according to the normal solutions of a differential equation of the fourth order. This equation is derived from that part of the distribution function which depends on the components of the relative motion. It is the equation obtained in attempting to find which stream functions ψ give a maximum value to the fraction

$$-\int_{-\frac{1}{2}}^{\frac{1}{2}} dy\, \overline{u'v'}\, \frac{d\lambda}{dy} \Big/ \frac{1}{R} \int_{-\frac{1}{2}}^{\frac{1}{2}} dy \overline{\zeta' v'}.$$

The complete genesis of the equation is given in § 5.

$N, \xi_1, \xi_2, \ldots, \xi_N$ We now postulate an N-dimensional phase space with coordinates $\xi_1, \xi_2, \ldots, \xi_N$. The amplitudes or coefficients with which the various normal solutions occur in the development are introduced as coordinates into the phase space, and thus serve as a framework for the construction of cells in this space. It will be seen later that a difficulty occurs as to whether one single enumeration of the modes of motion can serve both for the relative and the mean motions. For the present we shall discuss the relative motion in detail.

In his papers Burgers has put forward two different methods of developing a solution. In the first method he introduced a quadratic point lattice instead of the continuous xy coordinate field. The state of flow at any instant was then characterized by the values of the stream function at the finite number of lattice points. These values were then taken as coordinates in the phase space. Tollmien has undertaken the task of obtaining mathematically more correct results by making use of this phase space. He concludes that the velocity fluctuations with wave-lengths of the order of magnitude of the lattice spacing preponderate to an unallowable degree; and states that by the introduction of the point lattice a discontinuous element wholly foreign to the continuous fluid was introduced into the consideration. Tollmien[18,19] has suggested a method whereby the essential difficulties might be overcome, but it is not our purpose to examine this possibility here.

M In the second method Burgers assumes that the space between the parallel walls is observed at M instants, separated by equal intervals of time, M being a large number. It is further assumed that we are in possession of a classification of all possible modes of motion of the system under consideration, a classification which must satisfy the condition that all the enumerated modes of motion are of equal weight. These modes are taken to be of equal weight because it appears that the fields cannot be formed in various ways from other units. There is no simpler quantity which might be used to describe the probabilities of the various states of motion. The sequence of M observations made in the above way defines a microscopically described state of the system. The order in which the various modes occur in the sequence can have no influence upon the mean values which will be determined later. Hence if the modes are numbered 1, 2, 3, ..., the number of times—$n_1, n_2, n_3, \ldots$—each of them occurs in the sequence of observations uniquely determines a statistically described state. Now corresponding to one ensemble of fields of flow for the relative motion, there may be effected $M!$ permutations among

$n_1, n_2, n_3, \ldots$

the M observations. When those permutations, in which modes of motion of the same type change place, or differ only in a very slight degree, are not included, then there remain

$$W = M!/\Pi n! \qquad \text{.....(18)} \qquad W, \Pi$$

differently arranged sequences corresponding to the given set of numbers $n_1, n_2, n_3, \ldots$, where Π denotes the product of the n's.

Burgers's particular formulation and the fundamental assumption of the whole theory is to consider only quantities defined for instantaneously observed modes of motion, while the circumstance that the motion must have a stationary character over a long period of time is introduced explicitly as a special condition which must be satisfied by the sequence of quantities so defined. Of course it is really necessary, by an extra assumption, though according to Fowler[20,21] none has yet been given, to identify the long-time average so calculated with the short-time averages which are of physical significance. It is now assumed that the same methods may be used to calculate statistical mean values as are used in statistical mechanics as applied to the theory of gases. Now great importance is attached to the fact that the set of points in the multidimensional space of statistical mechanics satisfies Liouville's theorem. If τ is the density of a fine dust of representative points in any element of the phase space, this theorem may be stated in the form τ

$$\frac{\partial \tau}{\partial t} + \Sigma \frac{\partial}{\partial \xi_s}\left(\tau \frac{\partial \xi_s}{\partial t}\right) = 0 \qquad \text{.....(19),}$$

where the summation is for all elements of the phase space. The theorem shows that there can be no eventual crowding together of points into favoured regions of the phase space. The theorem in statistical mechanics is proved from the fact that the equations of motion are Hamiltonian. The systems in these cases are conservative whereas the system which we are discussing is dissipative. Now it has been shown in the dynamical theory of gases that for a dissipative system when left to itself the property does not hold. Burgers has given detailed consideration to the possibility that the set of points introduced in his theory satisfies Liouville's theorem. In his first formulation he only succeeded in establishing the result for an ideal fluid for which there can be no dissipation. Later it will appear, when the theory has been developed to a greater extent, that Liouville's theorem is not satisfied, but that a consideration of the theorem may lead to a restriction of the number of degrees of freedom to be considered in the statistics of turbulent motion.

§ 4. THE DISTRIBUTION FUNCTION

The equation (14) is used to construct the first condition which must be satisfied by a statistically described state of the system. If the motion shall be stationary it appears that the mean value of $\partial U/\partial t$ taken with respect to the time must eventually be zero, if the interval of time considered is long enough. This type of assumption is a familiar one in several branches of physics and its plausibility is quite evident. The right-hand side of equation (14) is now denoted by σ_m for the mode of motion σ_m

m numbered m in the enumeration of possible modes of motion constituting a state of the system. This quantity must satisfy the condition

Σ
$$\sum_m n_m \sigma_m = 0 \qquad \text{.....(20),}$$

together with
$$\sum_m n_m = M \qquad \text{.....(21).}$$

The condition (20) must be satisfied for all values of y, but it will not be necessary for it to be satisfied for $y = \pm\frac{1}{2}$, that is, at the walls.

The equation (16) must also be satisfied by the states of the system. The right-
ϵ_m hand side of equation (16) is denoted by ϵ_m for the mode of motion numbered m. Any state of the system must then satisfy the condition

$$\sum_m n_m \epsilon_m = 0 \qquad \text{.....(22),}$$

which condition applies to the whole field at once.

In order to avoid the use of the conception of a most probable distribution the lines indicated by Burgers in part III of his report are followed. Fowler's ([20], pp. 22–31) method of calculating averages by means of the introduction of a partition function is thus used. Average values are taken in preference to most probable values, which method has a great advantage in ease and rigour over mathematical presentation.

The statistical mean values of the n_m taken over a great number of sequences, each consisting of M observations, now have to be calculated. The weight of a statistically described state is given by W, and the statistical mean value of any one of the n's is then given by

$\bar{\ }, \Sigma^*$
$$\bar{n}_m = \sum_m{}^* W n_m \Big/ \sum_m{}^* W \qquad \text{.....(23),}$$

where the summation $\sum_m{}^*$ is to be extended over all the values of the n's consistent with the relations (20) to (22). If we integrate the relation (20) over the channel we can combine the relations (20) and (22) into the condition

$$\sum_m n_m \kappa_m = 0 \qquad \text{.....(24),}$$

κ_m where
$$\kappa_m = \epsilon_m - \int dy \lambda \sigma_m$$

λ and λ is a function of y.

From (18) and (23) we have

$$\bar{n}_m = \sum_m{}^* \frac{n_m M!}{\Pi n!} \Big/ \sum_m{}^* \frac{M!}{\Pi n!}.$$

Introducing the partition function

$\phi(z)$
$$\phi(z) = \sum_m z^{\kappa_m}$$

and applying the method of steepest descents ([20], p. 27), we find that

ϑ
$$\bar{n}_m = M \vartheta^{\kappa_m} / \phi(\vartheta) \qquad \text{.....(25),}$$

where ϑ is the unique positive fractional root of

$$\sum \kappa_m z^{\kappa_m - 1} = 0 \qquad \text{.....(26),}$$

since M is large. Since M and $\phi(\vartheta)$ are independent of m, we can conform to the usual notation by writing the result (25) in the form

$$\bar{n}_m = \exp(c + \beta\kappa_m) \quad \text{.....(27),}$$

c, β

where c and β are numerical constants and $\vartheta = e^{\beta}$. It may be noted that

e

$$\Sigma \bar{n}_m = M\Sigma\vartheta^{\kappa_m}/\phi(\vartheta) = M,$$

which correlates the relations obtained with the more obvious result (21).

Now from equations (14) and (20) we get, on integration by parts,

$$\int d\gamma\lambda\sigma_m = -\int d\gamma \left[\frac{d\lambda}{d\gamma}\left\{-\overline{u'v'} + R^{-1}\frac{\partial U}{\partial y}\right\}\right]_m - 2C\int\lambda\, d\gamma.$$

Hence from equations (16), (24) and (27) we obtain

$$\bar{n}_m = \exp\left\{c + \beta\int_{-\frac{1}{2}}^{\frac{1}{2}} d\gamma \left[\frac{1}{R}\frac{\partial U}{\partial y}\frac{d\lambda}{d\gamma} - \frac{1}{R}\left(\frac{\partial U}{\partial y}\right)^2 - \overline{u'v'}\frac{d\lambda}{d\gamma} - \frac{1}{R}\zeta^{\prime 2}\right]_m\right\} \quad \text{...(28),}$$

[ ]

where the index m indicates that the quantity between the square brackets must be calculated for the mode of motion numbered m, if we set

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} \lambda\, d\gamma = 1 \quad \text{.....(29),}$$

which device we see has eliminated the pressure-drop from the exponent in equation (27). The constants c and β are to be determined by the condition that equations (21) and (28) must be satisfied; the condition for λ is that equation (20) shall be fulfilled for all values of y satisfying the condition $-\frac{1}{2} < y < \frac{1}{2}$. Since equation (20) need not be satisfied at the walls we assume that λ is zero for $y = \pm\frac{1}{2}$.

These conditions which determine the constants c and β and the function λ occurring in the exponent of e in equation (28) can be expressed by introducing a function Φ defined by

Φ

$$\Phi = \sum_m \bar{n}_m,$$

where the summation with respect to m must be extended over all modes of motion, or in other words, over the whole of the system of cells into which the ξ-space has been divided. This summation may be replaced by an integration over the ξ-space. Then the function Φ must take the value M and at the same time it must be stationary with respect to arbitrary variations of β and of the function λ, the latter however being subject to equation (29) and to the condition that $\lambda = 0$ for $y = \pm\frac{1}{2}$.

§ 5. INTRODUCTION OF A SERIES OF NORMAL FUNCTIONS FOR THE DEVELOPMENT OF THE STREAM FUNCTION

Further progress with the theory can only be made by basing the method of enumerating the various modes of motion upon some special hypothesis. Burgers assumes that some particular value of λ has been chosen, and then investigates the distribution of the relative motion. This can be done since the exponent of equation

(28) can be divided into two distinct parts, one depending on the mean motion and the other upon the relative motion. For the moment then we consider the distribution function

$$\bar{n}_m = \exp\left\{\beta \int_{-\frac{1}{2}}^{\frac{1}{2}} dy \left[-\overline{u'v'}\frac{d\lambda}{dy} - \frac{1}{R}\overline{\zeta'^2}\right]_m\right\} \qquad \text{.....(30),}$$

and investigate which stream functions ψ give a maximum value to the fraction

$$-\int_{-\frac{1}{2}}^{\frac{1}{2}} dy\, \overline{u'v'}\frac{d\lambda}{dy} \Big/ \frac{1}{R}\int_{-\frac{1}{2}}^{\frac{1}{2}} dy\, \overline{\zeta'^2} \qquad \text{.....(31).}$$

To this purpose Burgers writes

$\chi_I, \chi_{II}, \alpha$

$$\psi = \chi_I(y)\cos\alpha x + \chi_{II}(y)\sin\alpha x \qquad \text{.....(32),}$$

and afterwards this result is generalized to include all possible combinations of stream functions. It may be however that types of relative motion built up from special systems of vortices, which would depend on the type of channel considered and the degree of turbulence, may predominate over other types. This possibility has been considered by Orr[22], Lorentz ([2], p. 48) and Burgers[17] but the investigations do not give much help here.

From (32) we deduce

$$-\overline{u'v'} = \frac{\alpha}{2}(\chi_I'\chi_{II} - \chi_I\chi_{II}'),$$

$$\overline{\zeta'^2} = \tfrac{1}{2}(\chi_I'' - \alpha^2\chi_I)^2 + \tfrac{1}{2}(\chi_{II}'' - \alpha^2\chi_{II})^2,$$

where dashes on the χ functions denote differentiation with respect to y, since

$$\frac{1}{L}\int_0^L \cos^2\alpha x\, dx = \frac{1}{L}\int_0^L \sin^2\alpha x\, dx = \tfrac{1}{2}$$

and

$$\frac{1}{L}\int_0^L \sin\alpha x\cos\alpha x\, dx = 0.$$

By using these values in the exponent of equation (30) we obtain

$$\beta\int_{-\frac{1}{2}}^{\frac{1}{2}} dy \left[\frac{\alpha}{2}(\chi_I'\chi_{II} - \chi_I\chi_{II}')\frac{d\lambda}{dy} - \frac{1}{2R}\left\{(\chi_I' - \alpha^2\chi_I)^2 + (\chi_{II}'' - \alpha^2\chi_{II})^2\right\}\right]_m \qquad \text{.....(33).}$$

Now,

$$\frac{\alpha}{2}\int_{-\frac{1}{2}}^{\frac{1}{2}} dy\,(\chi_I'\chi_{II} - \chi_I\chi_{II}')\frac{d\lambda}{dy} = \frac{\alpha}{2}\int_{-\frac{1}{2}}^{\frac{1}{2}} dy\,\chi_{II}^2\frac{d}{dy}\left(\frac{\chi_I}{\chi_{II}}\right)\frac{d\lambda}{dy}\,dy,$$

therefore, on integration by parts and use of the conditions

$$\chi_I = \chi_{II} = \chi_I' = \chi_{II}' = 0 \qquad \text{at } y = \pm\tfrac{1}{2} \qquad \text{.....(34),}$$

this expression becomes

$$-\alpha\int_{-\frac{1}{2}}^{\frac{1}{2}}\left(\chi_{II}'\frac{d\lambda}{dy} + \tfrac{1}{2}\chi_{II}\frac{d^2\lambda}{dy^2}\right)\chi_I\, dy.$$

Again repeated integration by parts and the use of (34) yields the result

$$\int_{-\frac{1}{2}}^{\frac{1}{2}}(\chi_I'' - \alpha^2\chi_I)^2\, dy = \int_{-\frac{1}{2}}^{\frac{1}{2}}(\chi_I^{iv} - 2\alpha^2\chi_I'' + \alpha^4\chi_I)\,\chi_I\, dy.$$

Collecting these results and using the corresponding ones for χ_{II} we eventually find that the maximum values of the fraction (31) are given by the two simultaneous linear equations:

$$\left.\begin{aligned}\chi_I^{iv}-2\alpha^2\chi_I''+\alpha^4\chi_I+\alpha R\Lambda\left(\frac{d\lambda}{dy}\chi_{II}'+\tfrac{1}{2}\frac{d^2\lambda}{dy^2}\chi_{II}\right)=0\\ \chi_{II}^{iv}-2\alpha^2\chi_{II}''+\alpha^4\chi_{II}-\alpha R\Lambda\left(\frac{d\lambda}{dy}\chi_I'+\tfrac{1}{2}\frac{d^2\lambda}{dy^2}\chi_I\right)=0\end{aligned}\right\}\quad\ldots\ldots(35),\qquad\Lambda$$

which may be combined into the single one,

$$\chi^{iv}-2\alpha^2\chi''+\alpha^4\chi-i\alpha R\Lambda\left(\frac{d\lambda}{dy}\chi'+\tfrac{1}{2}\frac{d^2\lambda}{dy^2}\chi\right)=0\qquad\ldots\ldots(36),$$

where $\chi=\chi_I+i\chi_{II}$. χ

The equation satisfied by the conjugate function $\bar{\chi}=\chi_I-i\chi_{II}$ is obtained by changing i into $-i$. By the rules for finding the maximum value of a function containing two arbitrary functions we can vary χ_I and χ_{II} separately. This accounts for the introduction of the number Λ, which is called the characteristic number; solutions satisfying the boundary conditions (34) can be obtained for a definite series of values of this parameter. If the corresponding values are denoted by Λ_k the corresponding normal solutions are written χ_k. $\bar{\chi}$ Λ_k χ_k

§ 6. APPLICATION OF THE NORMAL FUNCTIONS TO THE REDUCTION OF THE DISTRIBUTION FUNCTION

Certain properties of the solutions of equation (36) will now be developed. Let χ_k, χ_l be two functions belonging to different characteristic numbers Λ_k, Λ_l. If we multiply the equations satisfied by χ_k, $\bar{\chi}_l$ respectively by $\Lambda\bar{\chi}_l$ and $\Lambda_k\chi_k$ and subtract we get χ_l, Λ_l

$$i\alpha R\Lambda_k\Lambda_l\frac{d\lambda}{dy}(\chi_k'\bar{\chi}_l-\chi_k\bar{\chi}_l')=\Lambda_l\bar{\chi}_l(\chi_k^{iv}-2\alpha^2\chi_k''+\alpha^4\chi_k)-\Lambda_k\chi_k(\bar{\chi}_l^{iv}-2\alpha^2\bar{\chi}_l''+\alpha^4\bar{\chi}_l),$$

therefore on integrating by parts and using the boundary conditions (34) we obtain the relation

$$\int_{-\frac{1}{2}}^{\frac{1}{2}}dy\,\frac{d\lambda}{dy}(\chi_k'\bar{\chi}_l-\chi_k\bar{\chi}_l')=0\qquad\ldots\ldots(37a).$$

Six other relations can be proved in a similar way. These relations are:

that all the Λ_k are real $\ldots\ldots(37b)$,

$$\int_{-\frac{1}{2}}^{\frac{1}{2}}dy\,(\chi_k''-\alpha^2\chi_k)(\bar{\chi}_l''-\alpha^2\bar{\chi}_l)=0\qquad\ldots\ldots(37c),$$

$$\int_{-\frac{1}{2}}^{\frac{1}{2}}dy\,\frac{d\lambda}{dy}(\chi_k'\chi_l-\chi_k\chi_l')=0\qquad\ldots\ldots(37d),$$

$$\int_{-\frac{1}{2}}^{\frac{1}{2}}dy\,(\chi_k''-\alpha^2\chi_k)(\chi_l''-\alpha^2\chi_l)=0\qquad\ldots\ldots(37e).$$

To fix the arbitrary constant which occurs in any normal functions, we write

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} dy \frac{d\lambda}{dy} (\chi_k' \bar{\chi}_k - \chi_k \bar{\chi}_k') = -2i\alpha^2 R^{-1} \qquad \text{.....(37}f\text{)},$$

and then we find that

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} dy (\chi_k'' - \alpha^2 \chi_k)(\bar{\chi}_k'' - \alpha^2 \bar{\chi}_k) = \Lambda_k \alpha^3 \qquad \text{.....(37}g\text{)}.$$

These relations together with the equations obtained by changing every χ into its conjugate complex are valid for all values of k and l including the case when $k=l$.

Now with these normalizing conditions all the values of Λ will of necessity be positive, provided α is positive. It is convenient to retain the restriction of α to positive values and introduce complex quantities. Burgers thus introduces a complete stream function ψ_m for the mode of motion numbered m in the form

ψ_m

$$\psi_m = \tfrac{1}{2} \sum_{\alpha k} \{e^{-i\alpha x} (A_{\alpha k} \chi_{\alpha k} + B_{\alpha k} \bar{\chi}_{\alpha k}) + e^{i\alpha x} (\bar{A}_{\alpha k} \bar{\chi}_{\alpha k} + B_{\alpha k} \chi_{\alpha k})\} \qquad \text{.....(38)},$$

$A_{\alpha k}$, $\bar{A}_{\alpha k}$

where $A_{\alpha k} = A^{\mathrm{I}}{}_{\alpha k} + iA^{\mathrm{II}}{}_{\alpha k}$, $\bar{A}_{\alpha k} = A^{\mathrm{I}}{}_{\alpha k} - iA^{\mathrm{II}}{}_{\alpha k}$, etc.

$B_{\alpha k}$, $\bar{B}_{\alpha k}$

Any mode of relative motion is now specified by the values of the As and Bs, and in calculating statistical mean values the summation with respect to the number m can be replaced by an integration over the As and Bs. In the above expression the normal functions have been denoted by $\chi_{\alpha k}$ in order to denote their dependence on the parameter α. The summation with respect to α should really be replaced by an integration, but for simplicity of working the above form of the function ψ_m is retained.

The amplitudes with which these normal solutions occur are the coordinates introduced in § 3 into the phase space to serve as a framework for the construction of cells in this space.

From equation (38) we deduce that

$$-\overline{u'v'} = \frac{i}{4} \sum_{\alpha kl} \alpha [(A_{\alpha k} \bar{A}_{\alpha l} - \bar{B}_{\alpha k} B_{\alpha l})(\chi'_{\alpha k} \bar{\chi}_{\alpha l} - \chi_{\alpha k} \bar{\chi}'_{\alpha l}) + A_{\alpha k} \bar{B}_{\alpha l} (\chi'_{\alpha k} \chi_{\alpha l} - \chi_{\alpha k} \chi'_{\alpha l}) - \bar{A}_{\alpha k} B_{\alpha l} (\bar{\chi}'_{\alpha k} \bar{\chi}_{\alpha l} - \bar{\chi}_{\alpha k} \bar{\chi}'_{\alpha l})] \qquad \text{.....(39)},$$

$$\overline{\zeta'^2} = \tfrac{1}{2} \sum_{\alpha kl} [(A_{\alpha k} \bar{A}_{\alpha l} + \bar{B}_{\alpha k} B_{\alpha l})(\chi''_{\alpha k} - \alpha^2 \chi_{\alpha k})(\bar{\chi}''_{\alpha l} - \alpha^2 \bar{\chi}_{\alpha l}) + A_{\alpha k} \bar{B}_{\alpha l} (\chi''_{\alpha k} - \alpha^2 \chi_{\alpha k})(\chi''_{\alpha l} - \alpha^2 \chi_{\alpha l}) + \bar{A}_{\alpha k} B_{\alpha l} (\bar{\chi}''_{\alpha k} - \alpha^2 \bar{\chi}_{\alpha l})(\bar{\chi}''_{\alpha l} - \alpha^2 \bar{\chi}_{\alpha l})] \qquad (40),$$

from which the following expressions for the integrals can be obtained by making use of the relations (37),

$$-\int_{-\frac{1}{2}}^{\frac{1}{2}} dy \, \overline{u'v'} \frac{d\lambda}{dy} = \frac{1}{2R} \sum_{\alpha k} \alpha^3 (A_{\alpha k} \bar{A}_{\alpha k} - B_{\alpha k} \bar{B}_{\alpha k}) \qquad \text{.....(41)},$$

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} dy \, \overline{\zeta'^2} = \tfrac{1}{2} \sum_{\alpha k} \alpha^3 \Lambda_{\alpha k} (A_{\alpha k} \bar{A}_{\alpha k} + B_{\alpha k} \bar{B}_{\alpha k}) \qquad \text{.....(42)}.$$

Consequently the exponent in the distribution function (30) assumes the form

$$-\frac{\beta}{2R} \sum_{\alpha k} \alpha^3 \{A_{\alpha k} \bar{A}_{\alpha k} (\Lambda_{\alpha k} - 1) + B_{\alpha k} \bar{B}_{\alpha k} (\Lambda_{\alpha k} + 1)\} \qquad \text{.....(43)}.$$

This distribution function must remain finite for all values (finite and infinite) of the variables. To satisfy this condition all the characteristic numbers $\Lambda_{\alpha k}$ must be greater than unity. The statistical mean value $\overline{f}$ of any quantity f depending on the As and Bs can then be calculated by means of the formula

$$\overline{f}=(\Sigma \overline{n}_m f_m)/(\Sigma \overline{n}_m) \qquad \text{.....(44),}$$

where the summation with respect to m may be replaced by an integration over the As and Bs.

§ 7. INTRODUCTION OF A SPECIAL FUNCTION λ

Further development of the theory depends on the solution of the differential equation (36),

$$\chi^{\mathrm{IV}}-2\alpha^2\chi''+\alpha^4\chi-i\alpha R\Lambda\left(\frac{d\lambda}{dy}\chi'+\frac{1}{2}\frac{d^2\lambda}{dy^2}\chi\right)=0 \qquad \text{.....(36).}$$

The quantities Λ and λ determine the complete solution of the equation. The characteristic numbers Λ_k of the equation must be real and greater than unity. The function λ must satisfy the conditions $\int_{-\frac{1}{2}}^{\frac{1}{2}}\lambda dy=1$ and $\lambda=0$ for $y=\pm\frac{1}{2}$. In § 5, part v of his papers, Burgers provisionally assumes a plausible value of λ in order to arrive at an approximate result. We write $\xi=\alpha\,(y+\frac{1}{2})=\alpha\eta$ so that η is the distance from the wall $y=-\frac{1}{2}$. It is now assumed that there exists a certain similarity amongst the normal functions $\chi_{\alpha k}$. As the scale in the direction of the coordinate x is determined by the parameter α, it is possible that the functions $\chi_{\alpha k}$ might be functions of the single variable $\xi=\alpha\,(y+\frac{1}{2})$. As the presence of the other wall disturbs the similarity we must assume that the other wall is situated so far away that it does not upset the similarity. If then we find functions $\chi_{\alpha k}$ which decrease sufficiently quickly for large values of ξ, the presence of the other wall may be without appreciable influence upon them. Naturally it would be more desirable to discuss the normal solutions of equation (36) with an arbitrary form of $d\lambda/dy$, but the method of doing so has not yet been found.

Considering χ as a function of ξ, the equation (36) becomes

$$\frac{d^4\chi}{d\xi^4}-2\frac{d^2\chi}{d\xi^2}+\chi-iR\Lambda\left(\frac{1}{\alpha^2}\frac{d\lambda}{d\eta}\frac{d\chi}{d\xi}+\frac{1}{2\alpha^3}\frac{d^2\lambda}{d\eta^2}\chi\right)=0 \qquad \text{.....(45).}$$

If we assume that $d\lambda/d\eta$ is proportional to $1/\eta^2$, α and η will disappear as separate variables from this equation. The expression

$$\frac{d\lambda}{d\eta}=\frac{b}{R\eta^2} \qquad \text{.....(46),}$$

where b is a positive constant, is then applied to the neighbourhood of the wall $\eta=0$. But in order not to violate the condition $\lambda=0$ for $\eta=0$ the condition

$$\frac{d\lambda}{d\eta}=\frac{b}{R\delta^2} \qquad \text{.....(47),}$$

where δ is a small constant giving the thickness of the boundary layer, must be

assumed for the region $\eta < \delta$. A different treatment is thus introduced for the boundary layers near the walls. The condition (47) leads to the equation

$$\chi^{iv} - 2\chi'' + \chi - \frac{ib\Lambda}{(\alpha\delta)^2}\chi' = 0 \qquad \text{.....(48).}$$

To solve this equation we can use functions of the type $\exp(m\xi)$, then

B_ν, m_ν

$$\chi = \sum_{\nu=1 \text{ to } 4} B_\nu \exp(m_\nu \xi), \qquad \text{.....(49),}$$

where the m_ν are the roots of the equation

$$m^4 - 2m^2 + 1 - \frac{ib\Lambda}{(\alpha\delta)^2} m = 0 \qquad \text{.....(50).}$$

The condition (46) leads to the hypergeometric equation after making a transformation by Laplace's method. The equation to be satisfied may be written

$$\frac{d^2\omega}{d\zeta^2} + \frac{4\zeta}{\zeta^2-1}\frac{d\omega}{d\zeta} - \frac{ib\Lambda}{(\zeta^2-1)^2}\omega = 0 \qquad \text{.....(51),}$$

while we must have the values of

$$\exp(\xi\zeta)(\zeta^2-1)^2\left(\omega\xi - \frac{d\omega}{d\xi}\right) \qquad \text{.....(52)}$$

the same at both ends of the path of integration, the value of χ being given by

$\omega(\zeta), \zeta$

$$\chi = \xi \int d\zeta \exp(\xi\zeta)\,\omega(\zeta) \qquad \text{.....(53).}$$

By using the boundary conditions that $\chi, \chi', \chi'', \chi'''$ must be continuous at $\xi = \alpha\delta$, and that $\chi = \chi' = 0$ for $\xi = 0$, equations may be written down which if solved would give the six arbitrary constants which occur in the solution. After a long analysis in part v of his papers Burgers[10] finds that the system of normal functions is defined by the formula

$$\chi_{\alpha k} = N_k \alpha\eta \int d\zeta e^{\alpha\eta\zeta}\omega_k(\zeta) \qquad \text{.....(54),}$$

N_k, ω_k

where N_k is a numerical constant and ω_k is given by the equation

$$\omega_k = \left(\frac{\zeta+1}{\zeta-1}\right)^\gamma \left(\frac{2}{1-\zeta}\right)^3 F\left(1-2k,\ 2+2i\sqrt{k^2+k},\ 4,\ \frac{2}{1-\zeta}\right) \qquad \text{.....(55),}$$

γ, F

where $\gamma = -(k+1) + i\sqrt{k^2+k}$ and F denotes the usual hypergeometric function. The factor N_k must be determined from the condition (37f), where $d\lambda/d\eta$ is given by equation (46) down to $\eta = 0$; it is independent of α.

Using the properties of hypergeometric functions and Pochhammer's extension of the Eulerian integral of the first kind[23], it is found that the characteristic values of the parameter Λ_k are given by

$$b\Lambda_k = 4(2k+1)\sqrt{(k^2+k)} \qquad \text{.....(56),}$$

k

where k is a positive integer. The case $k = 0$ must be excluded as Λ_k must be greater than zero. These characteristic values are independent of α. The condition that all

Λ_1

the Λ_k shall be greater than unity requires $b < b\Lambda_1$, so that

$$b < 17 \text{ (approximately)} \qquad \text{.....(57).}$$

Λ_1 is considered since this quantity is the least of the Λ_ks. It will be seen later that the relation (57) leads to an approximate lower value to the resistance coefficient C.

It must be noted that in obtaining the above results Burgers finds it necessary to apply the conditions $\chi=\chi'=0$ for $\xi=0$ to the expression (53), when δ becomes vanishingly small. This procedure appears to violate the condition (47).

§ 8. CONSIDERATION OF THE MEAN MOTION

The general statistical principles set forth in the previous paragraphs are intended not only to lead to the distribution of the relative motion over a section of the channel, but also to the distribution of the mean motion. The point of view taken is that the statistical considerations should apply to the actual motion, built up of the sum of the relative and mean motions, which actual motion is described by a stream function Ψ. The phase space must now be extended to bring in the co- Ψ
ordinates describing the mean motion. Now the coordinates to be introduced into the generalized space in which the function Ψ is represented can be divided into two groups. One group, defining the ξ-space, is formed by the As and Bs used in describing the stream function ψ of the relative motion which constitute all those coordinates which give rise to zero mean values of Ψ (taken with respect to x). The second group must describe the distribution of the mean values $\overline{\Psi}$ as a function of $\overline{\Psi}$
y. Now the exponent of the distribution function (28) could be divided into two parts, one part depending on the relative motion, and the other on the mean motion; hence the second group of coordinates can be considered as being orthogonal to the coordinates of the first group. We thus investigate the distribution function

$$\exp \frac{\beta}{R}\int_{-\frac{1}{2}}^{\frac{1}{2}} dy \left\{\frac{\partial U}{\partial y}\frac{d\lambda}{dy}-\left(\frac{\partial U}{\partial y}\right)^2\right\} \qquad \text{.....(58).}$$

The statistical mean value of U has to be determined by means of this function. A certain difficulty presents itself in that the spacing of the coordinates which describes the course of the function $\overline{\Psi}$ should bear a relation to the spacing of the coordinates used to describe ψ. The exponent of equation (58) is not homogeneous with respect to U, and so does not lead to a variational problem from which a system of normal functions can be deduced.

Burgers in part VI of his papers tries a simple Fourier expansion of the type

$$\overline{\Psi}=\sum_{n=0}^{n=\infty} U_n \sin(2n+1)\pi y \qquad \text{.....(59).} \qquad U_n$$

In order that $\overline{\Psi}$ itself shall be $\pm\frac{1}{2}$ at the walls $y=\pm\frac{1}{2}$ we must have

$$\Sigma(-1)^n U_n=\tfrac{1}{2} \qquad \text{.....(60).}$$

Now we put $$\lambda=\Sigma\lambda_n\cos(2n+1)\pi y \qquad \text{.....(61),} \qquad \lambda_n$$

which expression makes $\lambda=0$ at the walls $y=\pm\frac{1}{2}$. Since the equation $\int_{-\frac{1}{2}}^{\frac{1}{2}}\lambda dy=1$ must be satisfied the condition

$$\Sigma(-1)^n\frac{\lambda_n}{2n+1}=\frac{\pi}{2} \qquad \text{.....(62)}$$

must hold. It should be noted that a general form of the function λ is used and not the special forms given by equations (46) and (47). If equations (59) and (61) are introduced into equation (58), and the exponent transformed into a sum of squares by putting

U_n'

$$U_n = \frac{\lambda_n}{2\pi(2n+1)} + U_n',$$

the statistical problem can be solved. The result is

$$\overline{U}_n = \frac{\lambda_n}{2\pi(2n+1)} + \frac{24(-1)^n}{\pi^4(2n+1)^4} \qquad \text{.....(63).}$$

Using the Fourier expansion for y this result leads to the formula

$$\frac{d\overline{U}}{dy} = \frac{1}{2}\frac{d\lambda}{dy} - 6y \qquad \text{.....(64),}$$

from which

$$\overline{U} = \tfrac{1}{2}\lambda - 3y^2 + \text{const.} \qquad \text{.....(65).}$$

Using the relations (46) and (47) and bringing in the fact that the value of λ must be continuous at $y = -\frac{1}{2} + \delta$ we eventually obtain the results

$$\overline{U} = \frac{b}{R\delta} - \frac{b}{2R(y+\frac{1}{2})} - 3y^2 + \tfrac{3}{4} \text{ for } \delta < y + \tfrac{1}{2} < \tfrac{1}{2} \qquad \text{.....(66),}$$

$$\overline{U} = \frac{b(y+\frac{1}{2})}{2R\delta^2} - 3y^2 + \tfrac{3}{4} \text{ for } 0 < y + \tfrac{1}{2} < \delta \qquad \text{.....(67).}$$

The result (64) leads to values which in the central part of the channel are much too high. The values of $\overline{U}$ depend on the values of b and δ which have to be determined.

In spite of this difficulty we proceed to consider the equation (20). This equation may be written

$$2C - \frac{d}{dy}\overline{u'v'} + \frac{1}{R}\frac{d^2\overline{U}}{dy^2} = 0.$$

Integrating with respect to y we obtain

$$2Cy - \overline{u'v'} + \frac{1}{R}\frac{d\overline{U}}{dy} = 0 \qquad \text{.....(68).}$$

Now let us assume the course of $\frac{d\lambda}{dy}$ described by

$<$ $0 < \eta < \delta$, $\qquad \lambda = \dfrac{b\eta}{R\delta^2}$,

$\delta < \eta < \frac{1}{2}$, $\qquad \lambda = \dfrac{2b}{R\delta} - \dfrac{b}{R\eta}$.

From the condition

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} \lambda dy = 1$$

$\cong$ we deduce that $\qquad \delta \cong 2b/R$ (approximately) $\qquad$ .....(69).

This result substituted in equation (66) leads to

$$\overline{U} \cong \tfrac{5}{4} - b/R \cong \tfrac{5}{4} \qquad \text{.....(70),}$$

when $y=0$, that is at the centre of the channel. This gives for the velocity in the axis 1·25 times the principal velocity V_0. In reality it is much less, being about 1·1 times the principal velocity[10]. Equation (69) shows that there is only one adaptable parameter in λ. As the pressure-drop $2C$ is still unknown, we can make equation (68) fit at two points at most, and then we shall obtain relations to determine b and C. Now when $\eta=0$ or $y=-\frac{1}{2}$, $-\overline{u'v'}=0$, since $\chi_{\alpha k}$ and $d\chi_{\alpha k}/d\eta$ are zero for $\eta=0$. Equation (64) when applied to the same point gives

$$\frac{1}{R}\frac{dU}{dy}\cong\frac{1}{8b},$$

where we have used the result $\delta\cong 2b/R$. From equation (68) we thus get

$$C\cong 1/8b \qquad \text{.....(71).}$$

The second point is taken at a distance from the wall $y=-\frac{1}{2}$ which is great compared with δ but small considered in comparison with the half-breadth of the channel. In this case the contribution of $d\overline{U}/dy$ to equation (68) may be neglected. As the term $2Cy$ in equation (68) may still be replaced by $-C$ we obtain

$$C\cong -\overline{u'v'} \text{ for } \tfrac{1}{2}>\eta>\delta \qquad \text{.....(72).}$$

The third equation involving C, b and β is got by considering the dissipation condition (20); written explicitly it becomes

$$C-\frac{1}{R}\int_0^{\frac{1}{2}}d\eta\,\overline{\left(\frac{dU}{dy}\right)^2}-\frac{1}{R}\int_0^{\frac{1}{2}}d\eta\,\overline{\zeta'^2}=0 \qquad \text{.....(73).}$$

Further progress can only be made by calculating $\int_0^{\frac{1}{2}}d\eta\,\overline{\zeta'^2}$. This integral, which will be investigated in § 9, can be expressed as a sum with respect to the indices k and α.

Now from equations (57) and (71) we deduce that

$$C>0{\cdot}0074 \qquad \text{.....(74),}$$

which is many times greater than the experimental values[24,25]. The lowest dependable value observed seems to be 0·0017. It is interesting to note however that this theory gives a value differing from zero, whereas the power laws make $C\to 0$ as $R\to\infty$. $\to\infty$

Again if the polynomial for ω_k is substituted into the integral (54), the integration of the separate terms leads to confluent hypergeometric functions. If the asymptotic expansion is used for very great values of $\alpha\eta$ we can obtain a very rough image of the function $\chi_{\alpha k}$. It is found that for great values of k the first term may be roughly approximated to by the following expression

$$\chi_{\alpha k}\cong c\,\exp\{-(\alpha\eta')^2/2k\}\,(\cos\alpha\eta'-i\sin\alpha\eta') \qquad \text{.....(75),}$$

where $\alpha\eta'$ is written for $\alpha\eta-k$ while c is a constant. From equation (32) the stream function of the mode of relative motion corresponding to $\chi_{\alpha k}$ is found to be η', c

$$\psi_{\alpha k}\cong c\,\exp\{-(\alpha\eta')^2/2k\}\cos(\alpha x+\alpha\eta+\text{const.}) \qquad \text{.....(76).}$$

This mode of motion consists of a system of skew vortices, having their centres on the line $\eta = k/\alpha$; the dimensions in the direction of x being $\pi/\alpha = \pi\eta/k$, while the dimensions in the direction of η are of the order $2\sqrt{(k/\alpha)} = 2\eta\sqrt{k}$.

If α is increased while k is kept constant, the dimensions of the system are diminished in proportion, whereas if for a constant α the value of k is increased, the system moves outward. Owing to the presence of the other wall at $\eta = 1$, this process cannot go on indefinitely. Thus it appears that k cannot increase beyond a certain amount of the order α. This point is of importance later.

§ 9. CALCULATION OF $-\overline{u'v'}$ AND $\overline{\zeta'^2}$ FOR LARGE VALUES OF η

The values of $-\overline{u'v'}$ and $\overline{\zeta'^2}$ are obtained from equations (39) and (40) if for $A_{\alpha k}\bar{A}_{\alpha l}$, etc. the statistical mean values are substituted. Now the system of normal functions have been calculated for a course of $d\lambda/dy$ described by equation (46) right down to $\eta = 0$. If we adhere to this system of functions and we wish to keep to the course of $d\lambda/dy$ given by equations (46) and (47), then the integrals (37a), (37f) and (37d) no longer have the values indicated.

However, if the values of η are great compared with δ, it may be assumed from the results obtained in § 8 that those functions $\chi_{\alpha k}$ which materially contribute to the values of $-\overline{u'v'}$ and $\overline{\zeta'^2}$ will be very small in the narrow region where $d\lambda/dy$ is given by equation (47). Hence for these functions the corrections to the integrals (37a), etc. may be neglected and the expression (43) for the distribution function retained. From this function we obtain the statistical mean values

$$\overline{A_{\alpha k}\bar{A}_{\alpha l}} = \frac{2R}{\beta\alpha^3(\Lambda_k - 1)}, \quad \overline{B_{\alpha k}\bar{B}_{\alpha k}} = \frac{2R}{\beta\alpha^3(\Lambda_k + 1)}, \quad \overline{A_{\alpha k}\bar{A}_{\alpha l}} = \overline{B_{\alpha k}\bar{B}_{\alpha l}} = \overline{A_{\alpha k}\bar{B}_{\alpha l}} = 0 \qquad \text{.....(77).}$$

Substituting these values in equation (39) we obtain

$$-\overline{u'v'} \cong R\beta^{-1} \sum_{\alpha k} i\,(\chi'_{\alpha k}\bar{\chi}_{\alpha k} - \chi_{\alpha k}\bar{\chi}'_{\alpha k})/\alpha^2(\Lambda_k{}^2 - 1) \qquad \text{.....(78).}$$

In order to carry out the summation with respect to α we can replace the summation by an integration. To do this the constant interval between the values of α in the sum is assumed to be given by a certain number θ. This assumption does not introduce a new arbitrary constant into the theory as θ always occurs combined with β as the product $\beta\theta$. Then we can write

θ

$$\sum_{\alpha}(\ldots\ldots) \cong 1/\theta \int_0^\infty d\alpha\,(\ldots\ldots).$$

By equation (54) we get

$f_k(\alpha\eta)$

$$\chi'_{\alpha k}\bar{\chi}_{\alpha k} - \chi_{\alpha k}\bar{\chi}'_{\alpha k} = \alpha b^{-1} f_k(\alpha\eta),$$

where

$$f_k(\alpha\eta) = bN_k{}^2(\alpha\eta)^2 \left\{ \int d\zeta\, \zeta\, e^{\alpha\eta\zeta} w_k \int d\zeta e^{\alpha\eta\zeta} \bar{w}_k - \int d\zeta\, \zeta\, e^{\alpha\eta\zeta} \bar{w}_k \int d\zeta e^{\alpha\eta\zeta} w_k \right\} \qquad \text{.....(79).}$$

The expression for $-\overline{u'v'}$ becomes

$$-\overline{u'v'} \cong R\,(b\beta\theta)^{-1} \sum_k (\Lambda_k{}^2 - 1)^{-1} \int_0^\infty \frac{f_k(\alpha\eta)}{\alpha}\, d\alpha.$$

This integral appears to be convergent and yields a number σ_k, independent of α. σ_k
Thus we find, if $\Lambda_k = p_k/b$, that p_k

$$-\overline{u'v'} \cong \frac{Rb}{\beta\theta} \sum \frac{\sigma_k}{p_k{}^2 - b^2} \text{ for } \tfrac{1}{2} > \eta > \delta \qquad \text{.....(80).}$$

A similar calculation starting from equation (40) leads to the result that $\overline{\zeta'^2}$ must be proportional to η^{-2}. As this result is true for $\eta > \delta$, it may be expected that in the integral

$$\int_0^{\frac{1}{2}} d\eta\, \overline{\zeta'^2} \qquad \text{.....(81)}$$

the upper limit may be replaced by ∞ without great error. Now the values of the integrals (37*c*), (37*f*), (37*e*) are not affected by the course assumed for the function λ. Hence equation (40) must remain true if $d\lambda/dy$ is determined by equations (46) and (47) taken together. If in equation (40) we substitute the approximate values of $\overline{A_{\alpha k} \bar{A}_{\alpha l}}$, etc. as given in equation (77) we find

$$\int_0^\infty d\eta\, \overline{\zeta'^2} \cong \frac{2R}{\beta} \sum_{\alpha k} \frac{\Lambda_k{}^2}{\Lambda_k{}^2 - 1} \cong \frac{2R}{\beta\theta} \int_0^\infty d\alpha \sum \frac{\Lambda_k{}^2}{\Lambda_k{}^2 - 1} \qquad \text{.....(82).}$$

This expression is divergent both with respect to k and with respect to α. The summation with respect to k probably may be limited, as shown at the end of § 8. The value of this integral then threatens to become infinite unless a limit can be found to the system of modes of motion.

§ 10. CONCLUSION

It will be seen from the above that Burgers's theory does not give many satisfactory results. The restriction of the motion to two dimensions is bound to make the numerical results grossly inaccurate. A further investigation of equations (8) and (11) may lead to better results, but the matter appears to be very difficult. The form in which the theory has been developed has no other justification than its apparent simplicity. Tollmien has suggested that the dissipation condition must be abandoned as the subsidiary relation required in the setting up of the turbulence statistics problem, since he considers that the use of this condition implies that viscosity plays a determining rôle, which appears not to be true in practice.

In part VII of his papers, Burgers has given a system of simultaneous differential equations which show properties analogous to those of the equations of fluid motion, in so far as the existence of solutions representing a state of turbulence is concerned. These equations are much simpler, and may serve as a means whereby the statistical methods can be tested. On the application of Liouville's theorem to this system it

appears possible that the number of degrees of freedom to be considered in the statistics of turbulent motion might be restricted. It might be possible to show that the integral in equation (82) would then be convergent. At present the matter is in a speculative state. The difficulty is to show that the statistical methods are applicable, and then discuss the possible modes of motion. It may be that the application of what are really microscopic methods to a macroscopic problem cannot be justified.

REFERENCES

(1) O. Reynolds. *Scientific Papers*, **2**, 575–7.
(2) H. A. Lorentz. *Abh. theoret. Phys.* **1**, 58–60.
(3) L. Prandtl. *Z. Ver. dtsch. Ing.* **77**, 105 (1933).
(4) T. von Karman. *Nachr. Ges. Wiss. Göttingen*, p. 58 (1930).
(5) G. I. Taylor. *Proc. roy. Soc.* A, **135**, 685 (1932).
(6) H. Bateman. *Report on Hydrodynamics*, Nat. Res. Counc., p. 335 (1932).
(7) T. von Karman. *Z. angew. Math. Mech.* **1**, 250 (1921).
(8) R. von Mises. *Z. angew. Math. Mech.* **1**, 425 (1921).
(9) F. Noether. *Z. angew. Math. Mech.* **11**, 224 (1931); **13**, 115 (1933).
(10) J. M. Burgers. *Proc. K. Akad. wet. Amst.* **32**, 414, 643, 818 (1929).
(11) —— *Proc. K. Akad. wet. Amst.* **36**, 276, 390, 487, 620 (1933).
(12) O. Reynolds. *Philos. Trans.* **186**, 123 (1895).
(13) C. W. Oseen, J. M. Burgers, L. Keller and A. Friedmann. *First Int. Cong. appl. Mech.*, Delft (1924).
(14) A. A. Isakson. *J. Soc. phys.-chim. russe*, **61**, 257 (1929).
(15) —— *J. Soc. phys.-chim. russe*, **62**, 313 (1930).
(16) L. F. Richardson. *Weather Prediction*, p. 94 (1922).
(17) J. M. Burgers. *Proc. K. Akad. wet. Amst.* **26**, 582 (1923).
(18) W. Tollmien. *Z. angew. Math. Mech.* **13**, 331 (1933).
(19) —— *Physics*, **4**, 289 (1933).
(20) R. H. Fowler. *Statistical Mechanics*, p. 8 (1929).
(21) G. D. Birkhoff and B. O. Koopmann. "Recent contributions to the Ergodic theory," *Proc. nat. Acad. Sci.*, Wash., **18**, 279 (1932).
(22) W. McF. Orr. *Proc. R. Irish Acad.* **27**, 124–8 (1907).
(23) Whittaker and Watson. *Modern Analysis*, § 12·43.
(24) Stanton and Pannell. *Philos. Trans.* **214**, 199 (1914).
(25) Jakob and Erk. *ForschArb. IngWes.* **1**, 267 (1924).

SOUND

COVERING THE YEARS 1932 AND 1933

By E. G. RICHARDSON, B.A., Ph.D., D.Sc.

§ 1. FREE PROPAGATION: SUPERSONICS

By far the greater part of the work which has been done recently on the propagation of sound in unrestricted media as distinct from bounded systems has been concerned with sources whose frequency lies above the audible pitch limit. For want of a better name we call these radiations *supersonics*, since, to speak of them as high-frequency sound waves, besides being tedious, savours rather of a *lucus a non lucendo*. The supersonic source has still been in almost every case the quartz oscillator, piezoelectrically excited, while the fertile method invented by Pierce a few years ago of making an opposing solid reflector face the quartz at a short distance and observing when that distance is an integral number of half-wavelengths, by the reaction on the electric circuit maintaining the oscillations, has formed the basis of many measurements of velocity of propagation. If the paradox may be excused, the author would describe the apparatus as a "Kundt's tube without the tube," since, with suitably sized quartz and holder, appreciable spreading of the sound does not occur and plane waves may be obtained without confining the gas within solid boundaries.

Dealing first with supersonics in gases, we find an increasing interest in the phenomenon first noted by Pierce, i.e. that certain gases absorb these radiations to a great extent. Carbon dioxide, for example, shows an absorption 40 times that given by the Kirchoff theory (as due to viscosity and heat-conduction) at a frequency of 10^5 c./sec.

Hand in hand with this anomalous absorption there goes a dispersion. Figure 1, after Kneser[1] and Groszmann[2, 3], shows how these factors vary in the region 10^4 to 10^6 c./sec. This phenomenon has been explained in terms of a delay in the change of transitional energy into vibrational energy of the molecules as the wave is propagated. When the time period of vibration becomes comparable with this relaxation time or mean life of a sound quantum, the molecules become stiffer to the vibration (to use the expressive phrase of Hubbard) with a consequent absorption of energy and rise in γ, the ratio of specific heats. Probably the simplest mathematical treatment is that of Henry[4], which follows.

Let E_x be the actual energy and E_T the energy which the molecules would have if in equilibrium at a temperature T, and suppose that

$$dE_x/dt = \beta^{-1}(E_T - E_x) \qquad \ldots\ldots(1),$$

where β is the period of relaxation of the vibrational energy. If the gas is subject to adiabatic variations of temperature of frequency $\omega/2\pi$,

$$T = T_0 + T_1 \exp(i\omega t),$$

$$E_T = E_0 + C_\omega T_1 \exp(i\omega t),$$

C_ω being the specific heat of vibration. Substituting in equation (1) and solving, we find

$$E_x = E_0 + C_\omega T_1 (1 + i\omega\beta)^{-1} . \exp(i\omega t).$$

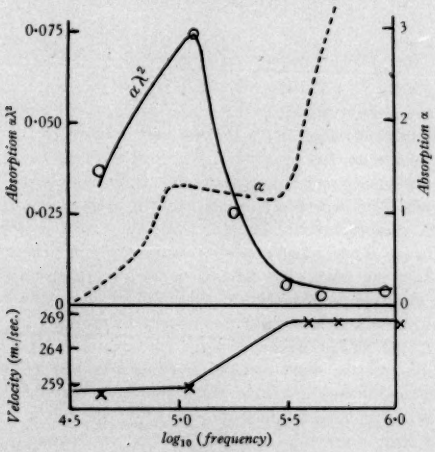


Figure 1.

If the total specific heat be regarded as made up of two parts, C_ω due to vibration and C_1 due to translation + rotation,

$$C' = C_1 + C_\omega (1 + i\omega\beta)^{-1},$$

where C' is complex; also, as usual, $\gamma' = 1 + R/C'$ and the velocity V of plane waves is $(\gamma' p/\rho)^{\frac{1}{2}}$, where ρ is the density and R is the gas constant.

Hence $\gamma' = 1 + R[C_1 + C_\omega/(1 + i\omega\beta)]^{-1}$ and is also complex. To find the real part we write:

$$[C_1 + C_\omega/(1 + i\omega\beta)]^{-1} = [1 + i\omega\beta]/[C_1 (1 + i\omega\beta) + C_\omega],$$

of which the real part is

$$\frac{C_0 + \omega^2\beta^2 C_1}{C_0^2 + \omega^2\beta^2 C_1^2} \qquad \ldots\ldots(2),$$

where $C_0 = C_\omega + C_1$, and is the total specific heat when ω becomes zero. Thus the usual specific heat at constant volume in the expression for the velocity of sound is replaced by the reciprocal of the expression (2). When ω approaches the value

$\beta^{-1}\, C_0/C_1$ a rise in velocity occurs from the normal V_0 to an ultimate value V_∞, where

$$V_0^2 = p/\rho.(1 + R/C_0),$$

$$V_\infty^2 = p/\rho.(1 + R/C_1),$$

p being the pressure of the gas. There is also absorption in the critical region due to a change of relative phase between pressure and condensation in the wave. From observations such as those of figure 1 it is then possible to predict the relaxation time. Thus for carbon dioxide at normal temperature and pressure $\beta = 6.10^{-7}$ sec. and for nitrous oxide[5] $\beta = 10^{-6}$ sec.

The theory promises to be of service to the physical chemist, since in a form originally worked out by Einstein in 1920, it may be applied to determine the degree of dissociation of a gas. In the above treatment we have considered a system in which the heat capacity work may vary with the frequency. With Einstein it was a question of the variation of the work of dissociation. In an actual dissociating gas, of course, both variations may intervene.

Richards and Reid[6] have studied the dissociation of nitrogen tetroxide in the equilibrium $N_2O_4 \rightleftharpoons 2NO_2$ by supersonic methods in the hope of being able to sort out the two variants. Both these lead to the same supersonic effects, viz. increase of velocity and absorption, the latter appearing as a change of phase between pressure and condensation. These writers, contrary to Teeter[7], think that it should be possible, at least in gases for which the heat of dissociation is large, to disentangle the two effects, particularly if the readings in the dispersion range be made at various pressures and temperatures. It should be realized that all these theories rest on equilibrium conditions and that while we have, in the simple outline above, spoken of the gas as a whole, the problem really relates to the probability that a given molecule shall be found in a given state. The more elaborate theories therefore follow a quantum rather than a macroscopic development such, for instance, as the recent attempt of Richards[8] to predict the behaviour of a gas with four or five independent energy states.

Kneser[9] has also attacked the case of a polyatomic gas, while Eucken and Becker[10] have found a marked effect of the admixture of certain gases (e.g. water and hydrochloric acid) on the transition from translational to vibrational energy in the carbon-dioxide molecule. Barnes[11], too, has examined mixtures of helium with other gases. Another theoretical refinement, due in the first instance to Luck[12], allows for departure from the ideal gas laws. Nevertheless, a mere physicist feels with Teeter that at present supersonics cannot compete in accuracy with the older methods such as spectroscopy, although it promises soon to become a useful handmaid in the investigation of molecular structure.

As regards experimental technique, the Pierce interferometer continues in use for velocity-measurements and the pressure on a torsion vane at various distances from the source, for absorption. Grabau[13] prefers a magnetostriction oscillator to the usual quartz and has paid particular attention to the irregularities always observed near the source. Avoiding this region, however, he opines that consistent values of the velocity of sound in air may be obtained. For absorption work,

Groszmann[13] and Pielemeier[14] prefer to catch the radiation upon a second quartz exactly similar to the source, placed at various distances from it. Some of Groszmann's results are shown on figure 1. This method, while an improvement on the torsion vane, will always be limited in range by the difficulty of getting pairs of crystals exactly matched in frequency. The traversing of a hot wire between source and stationary reflector, used by Richardson[15] and by Bücks and Müller[16], has the advantage of not requiring a disturbance of the maintaining system such as is unavoidable with the moving reflector of the usual Pierce apparatus. The resistance of the hot wire passes through a series of maxima (at the nodes) and minima (at the antinodes). The latter authors have also found that alcohol drops tend to gather into discs at the antinodes as in Kundt's tube at lower frequencies; see § 3. By either of these methods the half-wave-length in the gas can be measured.

Some discussion initiated by Hubbard[17, 18] has been carried on with regard to the possibility of using the original Pierce method to measure absorption as well as velocity. It is obvious that with an absorbent medium between source and reflector, the reaction on the former when the separation between the two is an integral number of half-wave-lengths will get less and less as this number increases; in fact in carbon dioxide at 200 kc./sec. the reaction disappears at a separation of a few centimetres. There is also a slight effect on the apparent wave-length as Bourgin[19] points out. The theory of the interferometer, including absorption, follows impedance lines (cf. § 4), but one doubts whether in practice an accurate value for the absorption coefficient can be obtained from it. But probably the same criticism applies to all the existing methods for measuring the absorption. The wave-length is much easier to measure with accuracy.

Some work of Knudsen[20, 21] also has a bearing on these problems, although it was actually done at audible frequencies. In examining the propagation of sound in rooms containing moist air, he found definite regions of selective absorption at low humidities (cf. figure 2) which Kneser explains in terms of collisions between the oxygen and water molecules, on the lines of the above theory for dissociating gases. It appears then that the new phenomena are not necessarily restricted to the supersonic range.

For liquids, the Pierce method continues in use. Swanson[22] has constructed an apparatus in which it is possible to measure supersonic velocities at pressures up to 20 atmospheres. This should prove useful for deducing the elasticity of liquids at high pressures from the sound-measurements. The heating produced in a liquid by the absorbed supersonic energy is considerable, noticeably when the volume is small. Thus Blancani and Dognan[23] found that 2 cm^3 of glycerine were raised 10° C. in 10 seconds by a moderate expenditure of power in the crystal source. It seems probable to the author that many of the biological effects observed on organisms contained in small water cells are due to this heating of the liquid. Certainly some control needs to be exercised in this direction in future experiments. There is also the question of the electro-magnetic radiations from the electrodes, which may reach the experimental culture. Malov[24] has used the heating effect to measure wave-length by traversing a *cold* wire resistance thermometer through the stationary waves

in oil between source and reflector, in the same way that the hot wire has been used in a gas. The wire is heated at the nodes, but the effect is usually much smaller than the corresponding cooling of a hot wire at the antinodes.

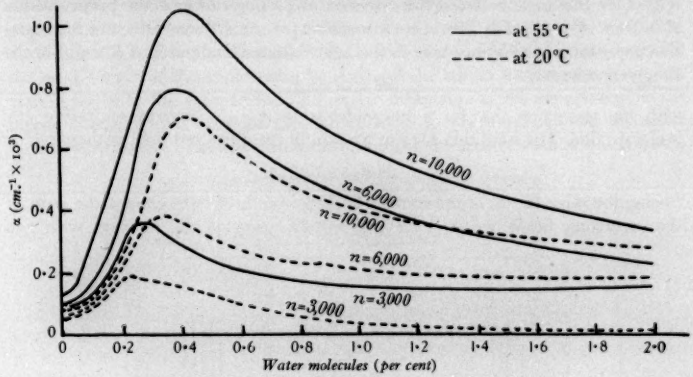


Figure 2.

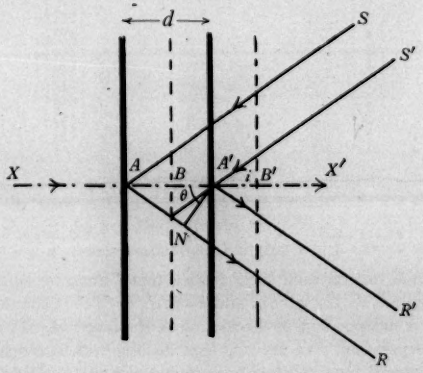


Figure 3.

About ten years ago Brillouin predicted that if a liquid were penetrated by progressive waves of compression of short wave-length and at the same time was irradiated by light, there would be diffraction of the light by the regular pattern of density-variations in the liquid in a fashion analogous to the diffraction of X rays by a crystal in the familiar Bragg experiment. Thus in figure 3, suppose that a sinusoidal supersonic wave is moving from left to right in the direction XX', giving at

the instant pictured maxima of compression distant d apart at A, A', with rarefactions at B, B', and let light of wave-length λ be incident at an angle i to the direction of propagation and be reflected along paths such as SAR, and let θ, equal to $\pi/2-i$, be the angle between the direction of propagation and the perpendicular $A'N$ from A' on to AR. Then the path-difference $=2AN=2d\sin\theta$ and the equation expressing equality of phase in the light reflected along AR, $A'R'$ is, as in the Bragg experiment,

$$m\lambda=2d\sin\theta$$

with the restriction that for a sinusoidal wave $m=\pm 1$, instead of having any integral value. The condition for a maximum in the diffracted light is therefore

$$\sin\theta=\pm\lambda/2l,$$

l being the wave-length of the sound. If then a beam of light crosses the path of the supersonic beam, we shall see the central image on the far side, with two

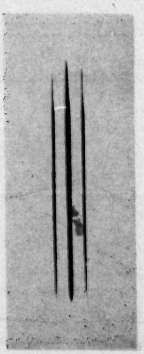


Figure 4 *a*.

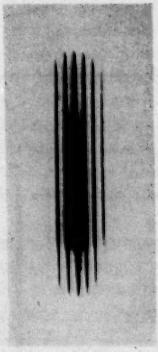


Figure 4 *b*.

diffraction maxima, one on each side. Notice that l must be small to produce a reasonable separation θ. If $\lambda=0{\cdot}6\mu$. and $l=10\mu$., $\theta=2°$. It should also be noted that the theory is independent of the motion of the layers A, A', as long as they conserve their separation. We are not here dealing with stationary waves. The necessary experiments have recently been undertaken on one hand by Debye and Sears [25] and on the other by Lucas and Biquard [26, 27, 28], who used light passing from a linear slit through a trough of liquid and out on to a photographic plate. Figure 4*a* shows one of the French photographs which is in accordance with theory, showing a central undeviated beam and a diffracted image on each side. Others, however, such as figure 4*b*, show three or five images. These have been ascribed to complexity in the supersonic wave-form, to excessive intensity of the supersonics, or to a sort of mirage produced by the successive layers of more dense liquid on the light; but the exact cause of these additional images remains uncertain at present.

In place of a slit, Bär and Edgar Meyer[29, 30] use an opaque screen studded with holes giving a large number of light-sources. Their plates (figure 5) have somewhat the appearance of Laue photographs, the dots being drawn out where the light has passed through irradiated liquid. Since the intensity of the diffracted light is proportional to the intensity of the supersonics at any point, it is possible to estimate the absorption suffered by the latter in the liquid. Bär and Meyer have also verified the wave-length of the supersonics by letting them fall on a wire grating set in a plane parallel to the wave-front. This causes diffraction of the supersonics in the liquid, changing their direction through an angle ϕ. There still remains the direct radiation, of course. The light from the section of the spotted source which passes

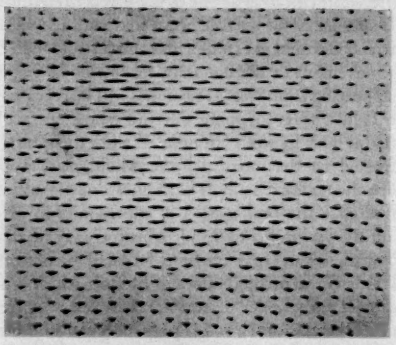


Figure 5.

through the diffracted beam *alone* will then have each spot drawn out in the same direction, viz. at an angle ϕ to the original beam. Measuring this angle and knowing the spacing of the wire grating they were able to calculate the supersonic wave-length, independently of a knowledge of the wave-length of the light.

All liquids show a greater absorption than that which theory would indicate as arising from viscosity and heat conduction, but some show an enormous absorption. Incidentally, the users of the Brillouin method seem able to use sources of very high frequency. Thus Bergmann[31] has been able to excite overtones of quartz oscillators up to the seventieth order and measure supersonic velocities in water, salt solutions and chloroform up to a frequency of 10^7 c./sec., corresponding to an electric wave-length of 12 m. Biquard[32, 33] gives the following figures for the absorption coefficients (α) at a frequency near 8×10^7 c./sec.:

Liquid	α calculated	α observed	Liquid	α calculated	α observed
Toluene	4·9	54	Ether	5·7	35
Xylene	7·8	47	Acetone	4·4	20
Ethyl acetate	6·7	49	Benzene	5·4	580
Methyl acetate	5·3	69	Chloroform	6·6	300

The enormous absorption in benzene and chloroform recalls the behaviour of carbon-dioxide and nitrous-oxide gases. It is indeed striking that those substances which behave abnormally to supersonic radiation are just those which scatter light most readily, and this leads one to speculate that there must be some connection between the mechanisms concerned. Possibly the common factor is the tendency to aggregation.

Turning to the propagation in solid bodies, we note that if these can be set directly into piezoelectric or magnetostrictive oscillation we can of course draw direct information from a study of the form of vibration of the oscillators themselves. In a posthumous paper, Dye[34,35] has published beautiful photographs of the appearance of quartz when vibrating as one of the plates in an optical interferometer. When the quartz is stationary the surface is covered with the usual pattern of interference bands such as one would get from a glass plate, but when the quartz vibrates a rapid oscillatory movement of the bands in the antinodal regions will occur, resulting in their disappearing from view. The nodes will be revealed by the unchanged appearance of the bands, thus mapping out the whole surface somewhat in the manner of the Chladni dust figures. Straubel[36] puts the quartz into a Michelson interferometer adjusted so that statically the whole plate appears dark. When vibration takes place the nodes remain as dark lines but the antinodes are lit up.

In order to investigate the propagation of supersonics in extended solids, it is necessary to pass on the vibrations to a rod. This rod may be directly stuck to one face of the quartz, or the crystal may transmit its oscillations to the rod through oil in a sealed vessel. This is the method adopted by Röhrich[37] who has investigated half-metre rods of steel and copper of various thicknesses at frequencies between 50 and 400 kc./sec. If the axis of the rod is perpendicular to the crystal-face, longitudinal vibrations are excited in it; if the axis is parallel to the face, transverse vibrations may be excited. Nodes in the rods were shown by lycopodium powder strewn on them. On the other hand, by suitably placed electrodes near one end of a nickel rod Giebe and Blechschmidt[38] have studied the propagation of longitudinal waves in nickel. A detector is located at a distance along the rod from the exciter and employs the reverse magnetostrictive effect. Both these investigations have as their main object the checking of a formula of Rayleigh according to which the frequency of a rod of circular section comparable in width with the wave-length is greater than that given by the simple formula

$$(2l)^{-1}\,(E/\rho)^{\frac{1}{2}}$$

in the ratio $1 : (1+\epsilon)$, where $\epsilon = \pi^2\sigma^2 j^2 r^2\,(2l)^{-2}$

and l is the length and r the radius of the rod, j the number of half-wave-lengths in the rod, E is Young's modulus, and σ is Poisson's ratio. It will be noticed that the departure from the simple formula increases with frequency, i.e. as j increases. There may in fact be stationary waves *across* the rod at certain high frequencies. The observed curves relating velocity with frequency are, as is predicted by theory, nearly parabolic except when resonance between the axial and radial waves occurs, and then a curve similar to those which occur in optics for anomalous dispersion is

obtained. The discrepancies between practice and theory are suitably accounted for by the two latter investigators, by invoking a coupling-coefficient between axial and radial oscillations.

Another interesting method of exciting the high-frequency longitudinal vibrations in rods is that of Kröncke[39] who succeeded in exciting longitudinal vibrations in rectangular rods as an edge tone by means of a jet impinging on the sharp edge of the material near the centre of the rod which was hung by two threads near its ends, dust showing the nodes.

Propagation in columns of fluid enclosed in tubes is reported upon in § 3.

The new tool has proved of service to the chemist in other directions than the pursuit of the elusive γ. Within the period covered by this report a number of investigators have tried the effect of supersonics on emulsions. Here again there is uncertainty in the control, for while some report that oil is readily emulsified in water by the associated turbulence, others find that a similar emulsion is broken down when put into a trough containing a piezoelectric oscillator.

On another aspect of this effect the chemists seem to be on surer ground, namely the breaking down of complex organic molecules into simpler ones. Szalay[40], using a 2-kilowatt supersonic oscillator at 723 kc./sec., has irradiated starch solutions and titrated samples with iodine from time to time. Whereas at the commencement the characteristic blue coloration of the mixture was obtained, after a period the transmitted light (as determined in a Pulfrich photometer provided with light-filters) was mainly in the red-brown region characteristic of dextrine, a simpler molecule. Viscosity-measurements made over the same period (30 minutes) showed a reduction of viscosity from that due to pure starch almost to that which a solution of pure dextrine of the same concentration should have. Gum arabic and gelatine also showed a reduction of viscosity, although it appears to the author that Szalay here rests his conclusions on uncertain grounds in view of the known anomalies in the viscosity of these jelly-like substances. Some aldehydes and ethers showed no change in physical or chemical properties after irradiation. Such investigations should be useful as showing the strength of the linkages in polymeric molecules.

§ 2. VIBRATING SYSTEMS: SOLIDS

There are two investigations on the pianoforte to be noticed. Mahajan[41] has been comparing the oscillations of different parts of the sound-board, to which mirrors were attached so as to yield graphic records of the vibrations. The acoustic spectra recorded at different points were very similar. His conclusion is that the sound-board vibrates as a whole, a result which is complimentary to the makers of the pianoforte in question, since it means that the sound will be radiated efficiently without interference from parts of the board vibrating out of phase. He also found that the soft pedal—which in his instrument interposed between hammers and strings the usual flannel flag beloved of upright piano manufacturers—produced a change of quality by rounding off the peaks in the recorded traces besides reducing

the amplitude. Presumably this change of quality would not occur on a grand piano where the pianissimo is effected by cutting out two of the three strings assigned to each note.

Sawada[42] has carried out some interesting comparisons between the standard and the electrical piano. First he found the best striking-position for the efficient production of sound in the wire. Expressing this efficiency as the ratio (energy of string) : (energy of hammer) he found that the best striking-position was near that at present adopted by the manufacturers (about one-seventh of the length from the end) but did not exactly coincide with the position for the best timbre (i.e. absence of overtones not consonant to the equi-tempered scale). In practice a compromise must be adopted in the choice of striking point. Whereas in the conventional instrument the vibrations of the strings are taken up and radiated by the sound-board, in the electric machine the vibrating strings are to act inductively on coils placed near them, and the consequent electric oscillations are amplified and passed on to a loud-speaker. In the former case the damping factor of the sound-board is not constant for all overtones; the interplay of rising and falling overtones constitutes that singing tone which is sought by connoisseurs. It can scarcely be imitated in the electric piano without sound-board, where the damping is more or less constant, so that the new instrument is likely to be out of favour with those accustomed to the old tone-quality. Sawada brings out the difference well in oscillographic records of the tones from the two instruments.

Two investigations relating to the tuning fork treat of the cause of the octave of the fundamental which may sometimes be heard, although it does not form one of the overtones of the fork proper. It is well known that sub-harmonics of the fork may be produced as stem tones when the fork is pressed on a bench. Yamashita and Aoki[43] establish by theory and experiment the fact that the octave may arise as a stem tone when the fork is screwed to a resonance box. In the theory the octave arises from the assumption that each point of the bars has a small vertical as well as a horizontal displacement. In the case examined by Derjaguin[44] the octave is impressed directly on the surrounding air through excessive amplitude of the fork. Thus in both cases it arises from non-compliance with the assumption of restoring forces proportional to the displacement in one plane, conditions recalling the Helmholtz theory of subjective combination tones.

Oliver[45], by an ingenious modification of Melde's experiment, has determined the velocity of transverse waves in paper. His apparatus is actually a system involving the coupled vibrations of two bars in resonance, one a steel reed maintained in transverse vibration, the other a stiff paper cantilever attached to one end of the reed. By adjustment of the length of the second or the frequency of the first, resonance and hence the wave-length of vibrations of known frequency in the paper, may be obtained. Presumably the work was undertaken in view of some industrial application, though this is not stated.

Zahradnicek and Zak[46] describe a circuit shown in figure 6, by which a valve may be applied to maintain a string in oscillation in any required mode. Two electromagnets are connected, one in the anode circuit and one in the grid circuit. In the

positions shown in the figure the fundamental will be excited, but by putting the electro-magnets opposite two neighbouring antinodes any other harmonic may be excited. For example, if the fifth harmonic is required the magnets are placed at one-tenth and three-tenths respectively, from one end. Such a device could no doubt be applied to the electric piano with an amplifier connected to the secondary of the transformer shown.

A considerable amount of work has been done by Colwell[47] in America and McLachlan[48] in this country on plates and conical diaphragms respectively, mainly on the theoretical side. In order to check his conclusions on the higher overtones of plates, the former investigator excites them by resonance with an audio-frequency oscillator and examines the Chladni figures formed on the plates. On the other hand, Flügge[49], with the same object in view, prefers to excite the plate through the maintained vibrations of a string attached to its centre in a sort of converse Melde experiment. Perhaps of greater interest are the investigations of Seiffert[50] (theoretical) and von Bekesy[51] (practical) into the vibrations of anisotropic plates, viz. those for

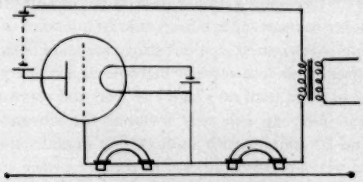


Figure 6.

which the velocity of transverse waves is different in directions which are perpendicular to one another; for anisotropy is a common property of wood, and a wooden sound-board is a feature of a number of musical instruments as well as of the partition walls of lightly constructed houses! In fact the velocity with the grain may be as much as three times that across the grain. The theory is worked out on Fermat's principle that the variation in the sum of the possible time-intervals by which the sound can pass from one point to another must be zero, but in the anisotropic plate equal times do not of course mean equal paths. It follows from the application of this and Huyghens' principle that as far as wave-propagation is concerned every anisotropic plate may be replaced by a wider isotropic one in a fiction which reminds one of conformal transformation. Von Bekesy has adduced an interesting ripple-tank demonstration in support of this theory. To imitate a cross-grained structure he sets up a number of submerged groynes made of parallel wooden laths 1 cm. apart, with their tops just below the surface. The waves from a point disturbance of the surface spread out in elliptical form in such a tank. He now surrounds two portions of the tank, one isotropic and the other anisotropic, with wooden boundaries to form triangular enclosures, and rocks the two wooden triangles up and down with the same frequency, until stationary waves, representing a mode of vibration of a triangular (isotropic) plate, are set up in the former region.

Before corresponding vibrations can be excited in the anisotropic region the base of the triangle (perpendicular to the laths) has to be reduced. Photographs are reproduced in the original paper from which one can compare the vibration-forms in the two cases.

In an attempt to produce a small acoustic measuring-instrument which shall not unduly disturb the field in which it is placed, Hall[52] has succeeded in constructing a tiny condenser microphone. The diaphragm consists of a sheet of aluminium foil 0·0004 in. thick and 2 cm. in diameter which is put under tension by screwing down a brass ring over the edge of a hard rubber bushing. The brass shell into which the system is screwed serves as the other plate of the condenser. Used as a transmitter the system has an output of about one millivolt per bar and is practically independent of frequency up to 6000 c./sec. Such an apparatus (and even smaller ones might be constructed) offers great possibilities for the plotting of sound-fields where larger microphones would distort the fields as in the experiments in diffraction of sound by plates described by Sivian and O'Neil[53]. The frequency range could also be raised, for Sacerdote[54] has made a similar instrument of aluminium 5 or 10 μ. thick and 0·8 cm. in diameter to respond to supersonic frequencies of 60,000 c./sec.

Mary Waller[55] has investigated an interesting extension of the instrument known as the Trevelyan rocker. This was actually discovered a century ago by Trevelyan's gardener, who had laid a hot fork on a block of lead and was surprised to find that it gave out a musical note. In this new instance the vibrations of tuning forks, plates, etc. are excited by contact with solid carbon dioxide, but the conditions for success are the same, viz. a good conductor in contact with a colder bad conductor of heat. If contact is uneven there will be some part of the plate which presses more heavily on the block than other parts. Transfer of heat will take place at this point with consequent melting of part of the block. In the case of a lead block expansion occurs which tips the rocker over and causes heavier contact of some other part, and so the process goes on, the low-frequency rocking producing a note. It is for this reason that a rocker provided with two ridges works best. With the solid carbon dioxide however another factor intervenes which makes for greater efficiency and allows of the maintenance by heat-transfer of the vibrations of objects of quite high natural frequency, which would be difficult to manage with the usual lead block. This factor is the sublimation of the carbon dioxide by the heat transfer, which sets up considerable gas pressures that help to tip the plate over. A high specific heat in the vibrator is naturally desirable, especially if its size is small, in order that the vibrations may continue for some time. Among the objects in which Miss Waller succeeded in exciting vibration were: tuning forks, bars, metal plates with Chladni sand figures, brass tubes, bells and small rings, piano wires, silver thimbles, table spoons, and quartz crystals. It was also found that ripples could be excited on the surface of mercury simply by dropping on it a solid piece of carbon dioxide of high specific gravity. In support of the rocker theory outlined above it was found that the ripples ceased when the under surface of the piece had its excrescences worn smooth, or when it had lost so much weight by sublimation that the contact pressure was insufficient.

Further work on bells has been done by Jones and Alderman[56] and by Curtiss and Giannini[57, 58]. The former have determined the acoustic spectrum of a bell at various epochs after striking. Since the bell is theoretically a struck anisotropic plate we find analogies with the struck string in the rate of decay of partials. The variation in this respect is one of the characteristics of bell tone. The latter workers have succeeded in examining the nodal lines by the method of Tyzzer, viz. excitation of each partial separately by means of a valve oscillator and loud-speaker, and exploration of the vibrating surface with a sort of stethoscope. The frequencies of the overtones, though inharmonic, are made to conform in a good bell to the ratios 1, 2, 2·4, 3, 4, 5, 6. Whereas the fifth partial is the strongest immediately after striking, it practically disappears after 1 sec., leaving the third partial master of the field. In 12 sec. this too has nearly fallen below audibility, leaving the fundamental which shows the least attenuation of all.

Table. Relative amplitudes of bell overtones at various epochs after striking. (Jones and Alderman.)

Order of partial	1	2	3	4	5	7
At strike	1	0·7	2·4	0·7	4	2·8
After ½ sec.	1	0·6	1·8	—	0·8	0·75
,, 1 sec.	1	0·6	1·8	—	0·1	0·4
,, 2 sec.	1	0·5	1·1	—	0·1	—
,, 3 sec.	1	0·4	1	—	—	—
,, 7 sec.	1	0·4	0·8	—	—	—
,, 12 sec.	1	0·2	0·4	—	—	—

This is for a tenor bell of 345 c./sec. In the treble the bells have a thicker waist (zone of minimum thickness of metal) and the attenuation of the partials is more nearly uniform. Giannini thinks however that the large bells with finer waists which have the acoustic properties outlined above have a better and more characteristic bell tone, and that this pattern, with a thick sound bow and thin waist, should be adopted throughout the peal. Carillons can further be improved by providing each bell with a damper, which is pulled off for a certain time as each key is pressed, like those of a piano. The time allowed for natural attenuation while the damper is off should not, however, be constant for all, but should be graded from treble to bass. A set of electrically operated dampers working on this principle has been successfully adapted to a carillon at Mountain Lake, Florida.

While we are on the subject of bells it is interesting to note that Kröncke[59] has successfully used the bell-like vibrations of small glass tubes to detect high-pitched sounds. A little tube of glass 2 mm. thick, sealed at both ends, 5 cm. in diameter, about 3 cm. long, and filled with dust is a typical example. When sound waves of frequency 17,800 c./sec. fall upon it, the glass vibrates with nodal lines parallel to the axis, the dust being whirled round inside. Note that the length is immaterial; this is not a case of longitudinal vibrations of the air as in Kundt's tube. Such detectors can be used in acoustic experiments wherever sensitive flames are used, and the motion of the dust can be exhibited to an audience*.

* They are marketed by Messrs Leybold, Berlin.

§ 3. VIBRATING SYSTEMS: FLUIDS

Dr Minnaert[60] is evidently one of those who find "sermons in stones, books in the running brooks, and good in everything." It is very pleasant in these electronic days to come across one of the fast vanishing race of natural philosophers who will carry out research on such a subject as "the murmur of the brook, the roar of the cataract and the humming of the sea, and endeavour to reduce such acoustical complexities to simpler elementary phenomena," to quote his own words. Now that the noise-meter is a standard apparatus in engineering methods it is indeed time that we should examine the sounds of nature in some detail. Nippoldt[61] has in fact been able to associate various degrees of turbulent weather with corresponding degrees of atmospheric loudness which he measures by the "masking effect." If this scale comes into common usage we may shortly see the weather charts issued by the meteorological office marked with loudness levels culminating in 70 decibels for a fierce storm! But to return to the work of Minnaert. His actual experiments consisted in allowing bubbles of air to form at the end of a tube beneath the surface of water in a beaker and rise to the surface. The pitch of the associated tone was determined by ear. The characteristic plop of the air bubble is not formed, as most people think, when it bursts at the surface (although a slight sound occurs then also) but at the moment of closure at the end of the tube. Compressional vibrations of the air in its envelope are formed, Minnaert thinks, by radial extensions and contractions of the envelope. With this type of displacement he shows that the square of the time period of a bubble of radius r, in a *liquid medium* of density ρ, is

$$4\pi^2\rho r^2\,(3\gamma p)^{-1} \quad \text{or} \quad n_1=(2\pi r)^{-1}\,(3\gamma p/\rho)^{\frac{1}{2}},$$

whereas a spherical resonator with rigid envelope would have a frequency (according to Rayleigh) of

$$n_2=0{\cdot}715r^{-1}\,(\gamma p\sigma)^{\frac{1}{2}},$$

σ being the density of the *gaseous medium* inside the envelope. This makes n_2 equal to $80\times n_1$ approximately, in fact, in most cases n_2 would be supersonic; moreover it would be raised considerably as the temperature raises the velocity of sound in air. The formula for n_1 has been compared with experiment and good agreement has been found. The frequency is independent of temperature and the nature of the gas (hydrogen gives the same results as air), but it is inversely as the radius of the bubble. A third possibility, capillary vibrations of the envelope, was ruled out on the ground that a change in the surface tension of the liquid did not alter the pitch.

Wind-tones continue to furnish a fertile subject of investigation in the domain of fluid vibrations, thanks to our increased knowledge of the physical background of hydrodynamics. The Abbé Carrière[62], whose work in sound is now well known, has investigated the edge-tone system formed when air from a linear slit impinges on the corner of the *internal* angle formed by two planes intersecting. The fluid in escaping from the corner flows down the pent sides and forms alternate eddies as it disengages from each plane in turn. Black and white pictures of these eddies are

given, in which it is apparent that it is the height of the dihedron above the orifice of the jet which determines the wave-length of the eddy system as one would expect, but the width of the obstacle also has an influence on the frequency. When the obstacle was freely suspended, resonance sometimes occurred between the edge tone and the suspended system, causing it to swing to and fro, a feature which is sometimes noticed with the conventional edge-tone apparatus.

Volkovsky[63] has observed that a plate dipped in Chinese ink and held edgewise to a jet whose velocity exceeds that of sound has a reiterative pattern impressed upon it when the distance from edge to slit has certain values. The pattern, which shows bulges, looking something like stationary waves, at equal intervals where the ink has been pushed aside, evidently is intimately related to the compressions and rarefactions in the space between the jet and the edge. We avoid the term "vortices" here advisedly since we are dealing with the region of supersonic velocities where the compressibility of the gas must intervene (cf. Hartmann's jet oscillator), but the phenomenon is evidently linked with the formation of stationary waves between jet-slit and edge, and we are justified in classifying it among edge tones.

Thom[64], in dealing with the slow motion of a cylinder through a fluid, gives some good photographs of the early stages in the formation of alternate vortices in the wake, to which aeolian tones are due. Also in this connection should be noted a (theoretical) paper by Marty[65] on the lines of flow caused by a plate *oscillating* in a fluid. There are three investigations on jets in relation to sound waves to be recorded. Tyler[66] has verified that the same criterion holds for frequency of drop-formation in a liquid jet as for vortex formation in wind-tones; that is to say, V/nD is constant where in this case V is the velocity of efflux of the jet from the nozzle, n is the frequency and D is the diameter of the nozzle. The jet is alternately broken up by swellings and contractions which are formed by casual disturbances on its surface near the nozzle and grow as the jet falls. Rayleigh has shown that the wave-length of such disturbances is more efficacious when it is equal to a certain multiple, 4·51, of the diameter; or in more precise words, of all disturbances which may be inflicted on the exuding fluid, the one with this wave-length will break up the jet before the others get a chance of doing so. If then we record the passage of the drops after formation past a fixed point we can record their frequency. This will be equal to $v/4{\cdot}51d$, where v is the velocity and d is the diameter of the jet at the break-away. If we assume that $d \propto D$ and $v \propto V$, we immediately derive the equation $V/nD = \text{constant}$. It is in this way, for instance, that a trickle of water falling from the scullery tap on to a plate in the sink often gives out a musical note. Dr Tyler's method of recording the frequency was to allow the drops to interrupt a beam of light falling on a photo-electric cell, the intermittent current in which was amplified and applied to a loud-speaker. Incidentally, we may note that the frequency so observed will be that to which the jet is most sensitive. This remark leads us on to the other two papers under this heading, which are on sensitive flames. It is not so much the ignition which divides these from sensitive jets as the absence, more or less complete, of surface tension in the gas—gas interface. Both G. B. Brown[67] and Zickendraht[68] used a valve oscillator for the generation of the tones to which the

flames were subjected, though the former provided greater safeguards for the purity of the tone. Photographs of the flame taken both under long exposure (to show its visual appearance) and instantaneously are given in the respective papers. Dr Brown used a large number of different jets. He detected a characteristic frequency for each jet (see above) by the interesting phenomenon of beats between the jet and the oscillator. This characteristic frequency could readily be found by tuning out the beat, in the same way that one tunes a wireless receiver. Symmetrical varicosities (to use the late Lord Rayleigh's expressive term) were observed to travel up such jets in the same way that they travel down water jets. More often, particularly with smoke jets, the vagarious motion is not symmetrical, being at its maximum in a plane passing through the jet and the source of sound. Seen from a direction perpendicular to this plane, the jet would often appear undisturbed. The sinuosities appear to start from a point a little above the orifice and fork outwards from this point with regularly increasing amplitude, very much as edge tones do. This is when the jet is responding: when quiet, it rises in a steady column to a considerable height. To be responsive, the velocity of efflux must be such that the flow is nearly, but not quite, turbulent. Sound of the right frequency then acts as a trigger which spills the flow over the critical velocity. As a matter of fact, it appears that ignition of the jet is not an improvement, since the outer layer of hot gas is insensitive to sound.

The two new facts which Zickendraht has discovered are (i) that the flames are almost equally sensitive to electrostatic and mechanical disturbances, as he found by removing the loud-speaker from his oscillator; and (ii) that, as in the case of the singing flame, stationary waves are formed in the supply tube of the flame. The length of this tube probably has an influence on the natural frequencies of the jet.

Turning now to the phenomena in the interior of a sounding pipe we have first to notice an extensive research by Andrade[69, 70] on the vibration in a Kundt's tube maintained at one end by a telephone diaphragm to which a simple harmonic vibration is given through a valve oscillator. The fact that overtones are practically absent, and the possibility of using considerable power, give advantages to this method of excitation over the old rod and resined cloth. In addition, the frequency is of course continuously variable. The use of considerable intensity and the mixture of tobacco smoke with the air enable very clear photographs to be made of the circulations which run from node to antinode, and which Rayleigh had predicted (figure 7). At a frequency of 600 c./sec. and a double amplitude of 0·48 mm. the rate of flow in the circulation of the smoke was 2·7 mm./sec. Another secret of the success and regularity of these pictures was the minimizing of convection by water-jacketing the tube. The observed stream lines agree with the theoretical ones for the velocity potential ψ, of which Rayleigh's formula is

$$\psi = C\,(r^4 - r^2R^2)\sin 2kx$$

at a radial distance r in a tube of radius R, and an axial distance x in a wave-length $2\pi/k$.

If a cylindrical obstacle is placed diametrically to the tube at an antinode and if the intensity is lowered somewhat, a striking phenomenon is observed. The cir-

culation is interrupted at the obstacle and a vortex system consisting, when seen in section, of four large vortices extending over a considerable portion of the tube makes its appearance. Similar effects were obtained with spherical obstacles, but in each case the vortex system became less distinct if the obstacle was not at an antinode. It seems also from the photographs that the circulation became disturbed by the asymmetrically displaced obstacle, the currents transverse to the tube shifting to the place of insertion of the obstacle. The theory of such vortex

Figure 7.

Figure 8.

systems has been worked out under the title "Calculation of plane periodic boundary-layer flow" by Schlichting[71]. He has calculated the stream function for a cylinder executing a simple-harmonic motion of small amplitude (i) in an infinite fluid, and (ii) in a fluid bounded by a concentric cylinder. In the first case, fluid flows in towards the cylinder from directions normal to the direction of oscillation. In the second case the ends of the four families of stream lines thus formed join up round the walls of the containing vessel to form four big vortices. Figure 8 shows one quarter of the field in this case. Schlichting's own attempts to obtain such vortices with a cylinder oscillating in still water failed owing to the large value of Reynolds'

number employed and consequent smallness of the vortex system, but he succeeded in correlating his theory with measurements of Andrade's photographs, which correspond, of course, more closely to the second case. The boundary of the containing tube or vessel is therefore an important feature in the formation of these vortices. The author emphasizes this fact because there is another vortex phenomenon when an obstacle is exposed to an alternating draught; this was first noticed by Carrière. His vortices are smaller in diameter than the cylinder, hugging it very closely, and are probably formed in the boundary layer itself (cf. shaded portion of figure 9). In this respect they resemble those pictured by Thom [64] and discovered by Porter, which lie behind an obstacle in one-way streaming at low Reynolds' number, whereas those of Andrade are much larger than the obstacle, they are

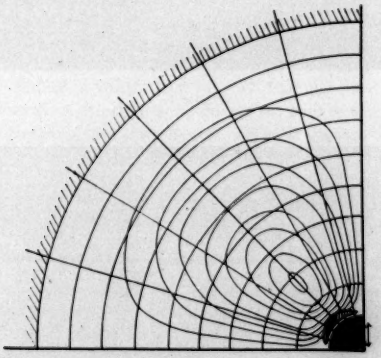


Figure 9.

bounded by the confines of the tube; and perhaps, on account of their size, they are better described as circulations than as vortices. Both types are covered by Schlichting's theory.

Prof. Andrade has also explained the formation of ridges in a Kundt's tube in terms of such vortices formed round dust particles. The ordinary dust employed is too heterogeneous to show the mode of formation well, but when a few graded wax spheres 0·29 mm. in diameter are put into the vibrating column and their individual vortex systems made visible by smoke, it can be seen that these systems cause two particles whose line of centres is parallel to the axis of the tube to maintain their separation unless they are sufficiently close together for the systems to coalesce. When coalescence occurs it is accompanied by a movement of the particles to a side-by-side orientation across the tube. When a number are in the tube they then range themselves in ridges whose longitudinal spacing varies from node to antinode with the amplitude (the size of the vortex system depending on this factor too) and the edges of the corresponding vortex systems are then contiguous. In fact, if there

are smaller particles in the tube, these are often swept into subsidiary ridges halfway between the main ones by the upward current of the vortex systems where these coalesce. Another striking feature of the maintained Kundt's tube when considerable power is used is the sweeping-up of the dust in the main circulations into discs at the antinodes; cf. figure 8, which is a photograph of part of the tube near an antinode under these conditions, and shows an antinodal disc and a number of neighbouring ridges.

Carrière[72] also has examined the motion of individual smoke particles in a sounding tube by compounding them with a perpendicular motion to give Lissajous' figures. The compounding vibration was obtained from an alternating electrostatic field normal to the axis of the tube, and fed by the same oscillator which excites the acoustic vibrations. The particles became electrically charged in passing through the pulverizer before being admitted to the tube. Ellipses or figures of eight were in general obtained, depending on the relative phase of electric and acoustic field at the place of observation. Rilbentrop[73] has revived the mechanical-valve method of measuring the amplitude in a pipe. As a matter of fact, his rectifier valve is placed at one end of the pipe, in the piston, while sound is fed into the other end from a loud-speaker. The valve opens in synchronism with the compressions, so that a steady pressure can be produced in a manometer. Curves showing how the pressure varies with the position of the piston are given.

There has not been much work on wind instruments during the period under revision, but Auger[74] has investigated the behaviour of a pipe of the clarinet stop of a reed organ from the point of view of the variation of frequency with pressure. When the coupled system comprising the reed and the pipe is functioning most efficiently, the motion of the reed is sinusoidal. At certain pressures, characterized by jerky movements of the reed, a raucous unmusical tone is produced. No theoretical explanation is given. Obata and Tesima[75] have examined the Japanese instrument syakuhati which is blown like a flute but has no mouth hole, so that the player has to blow across the open end in the same way that one blows a door key. Unlike that of the flute, the tone is rich in harmonics (why, is not clear) and these interfere in certain directions round the player so that the quality is markedly dependent on the relative positions of player and listener. The Japanese scientists have plotted out, by taking oscillograph records in conjunction with a mechanical playing mechanism, the sound-fields of the various harmonics. The tone is at its richest in the vertical plane through the instrument, supported horizontally in this case. Rather similar are the results of Sato[76,77] on a conical horn as collector of complex sound for reproduction purposes. He finds that the source must be directly in front, otherwise certain overtones will interfere and be missing in the reproduction.

Finally, there are three important researches on the velocity of sound in tubes. Kaye and Sherratt[78] use a tuned-tube resonator with the object of verifying the Kirchoff formula for the tube correction to the velocity of sound. Six different gases in tubes of six different diameters and of glass, copper or carbon were used to provide a wide range of fluid viscosity and boundary roughness, thermal con-

ductivity, etc. Resonance was detected by the reaction on the supply oscillator, which was adjusted to give a series of frequencies varying from 500 to 27,000 c./sec. For the lower frequencies a telephone diaphragm was used as transformer of electric into acoustic vibrations; and a quartz oscillator was used for the upper frequencies. The observations of velocity V at two temperatures 0° and 100° C. were reduced to free-space values V_0 by the formula originally due to Helmholtz but modified by Kirchoff:

$$V = V_0\{1 - c\,(2r)^{-1}(\pi n)^{-\frac{1}{2}}\},$$

where

$$c = \mu^{\frac{1}{2}} + (\sigma/\gamma)^{\frac{1}{2}}(\gamma - 1)$$

and μ is the coefficient of viscosity and σ the thermal diffusivity of the gas. It is satisfactory to find that, as a result of their careful measurements, the authors are able to vindicate this formula, at least for diameters down to 1 cm., for the question of its validity has aroused much controversy in the past. Some slight variation with the roughness of the tube-surface was found, and this, of course, the formula does not pretend to deal with.

Vance[79] proposes a modified Helmholtz equation to suit his recent measurements in air contained in glass tubes 0·1 to 3 cm. wide with supersonic frequencies of 30 to 200 kc./sec. Below $\frac{1}{2}$ cm. there is a very big change of velocity with diameter. Van Itterbeek[80] has made some resonator measurements of the velocity in gases at low temperatures. The resonator was of small size, 30 cm. × 2·6 cm., and so could be cooled down to temperatures of a few degrees absolute. The source was a transmitter at the bottom, while the resonant length was found by an adjustable stethoscope. Measurements at audible frequencies in helium, oxygen and hydrogen down to 20° K. and 0·1 atmosphere have been reported with a view to obtaining values of the velocity and its variation with pressure at low temperatures. The Helmholtz tube correction was applied, but in later work this author promises to use a much wider tube, when the correction will be negligible. Canac[81] reports measurements of the absorption of sound in tubes of elastic material, namely rubber and flexible metal. The intensity at different points along the tubes (2 cm. down to a few mm. wide) was compared aurally with that at different points along the stationary waves in a secondary but wider tube resonator, in which therefore the intensity-variation could be calculated if the motion was assumed to be simple-harmonic. Naturally the absorption is greater in the rubber tube, and increases rapidly with diminishing diameter. Further it increases with frequency, the observations going up to 1000 c./sec. Whether stationary waves were observed in the experimental tubes, or how these affected the observations, is not stated. In this connection Rocard[82] has modified the equations for sound-propagation in resonators to allow for absorption of this type, due to flexible walls, but the equations remain at present unsolved.

§ 4. ACOUSTIC IMPEDANCE AND PROPERTIES OF MATERIALS

Impedance methods are now well established in theoretical problems concerning the transmission of sound, but there is still a paucity of methods for measuring this quality. The difficulty in making an absolute measurement arises from the fact that

values of both alternating velocity and alternating pressure must be obtained, and although it is true that one can make a comparison between one impedance and another by using pressure measurements, for instance, each impedance is necessarily modified by that to which it is connected, and the expressions "known impedance" and "unknown impedance" in such comparisons may involve some self-deception. Possibly one day we may all agree upon an experimental standard impedance, but at present each investigator adopts his own, which is standard as far as his apparatus is concerned. Flanders[83] has devised a useful impedance-meter in which the "known" impedance Z_0 is a tube one-eighth of a wave long, which is joined in turn with the "unknown" impedance, to a tube of fixed length, the r.m.s. pressure at the junction being read on a condenser microphone. The apparatus, figure 10, has the advantage that the ultimate readings are made on a potentiometer. Sound from the

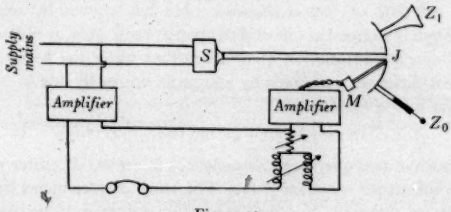


Figure 10.

loud-speaker S passes along the fixed tube to the junction J, where a side branch leads to the microphone M. The electric amplitudes of transmitter and receiver after amplification are compared on a potentiometer consisting of a variable resistance and a variable mutual inductance in series. The voltage across the output amplifier is balanced by a null method against that across the resistance and secondary of the inductance. In the analogous electrical case an electromotive force E is in series with an impedance T corresponding to the impedance up to the junction (or more precisely, with the impedance at the junction looking towards the same when on open circuit, the latter proviso implying a rigid stop at J in the acoustic case) and an attached impedance beyond J. The pressure E at J is, like the potential difference between the terminals at J and the other end of the cell, equal to $EZ(T+Z)^{-1}$.

We first attach the known impedance Z_0, next a rigid stop, and finally the unknown impedance Z_1, and we have:

$$Z=Z_0, \quad e_1=EZ_0(T+Z_0)^{-1},$$
$$Z=\infty, \quad e_2=E,$$
$$Z=Z_1, \quad e_3=EZ_1(T+Z_1)^{-1}.$$

The method will recall to the mind of the reader the well-known experiment of the elementary physics course, of finding the internal resistance of a cell.

Eliminating E and T we have

$$Z_1/Z_0=(e_2/e_1-1)/(e_2/e_3-1).$$

The pressures e are proportional to the respective currents through the resistance and primary. The drop in voltage across the resistance and secondary is equal and opposite in phase to E, when no current passes through the headphones. If then z is the (electrical) impedance value of the corresponding resistance r and inductance m, $e \times z$ is constant, and the equation becomes

$$Z_1/Z_0 = (z_1/z_2 - 1)/(z_3/z_2 - 1) = (z_1 - z_2)/(z_3 - z_2)$$
$$= [r_1 - r_2 + i\omega\,(m_1 - m_2)]/[r_3 - r_2 + i\omega\,(m_3 - m_2)].$$

Impedances of tubes, horns and orifices, determined in this apparatus at various frequencies, are exhibited by means of graphs in the paper.

Bate[84] has determined the resonant frequencies of a system which may be regarded as consisting of two pipe impedances in series, the pipes having different diameters D and d and lengths L and l. This is, in fact, a stop used in the organ under the names of *röhrflöte* or *flute-à-cheminée*. He has shown by experiment that whether the system is stopped at one end or open at both ends, as is usual, or whether the narrow pipe is a continuation of or re-entrant upon the wider, the resonant frequencies are sufficiently well given by a formula originally due to Aldis which we prefer to express in the form

$$i\rho c \tan(2\pi l/\lambda)\, d^{-2} + i\rho c \tan(2\pi L/\lambda)\, D^{-2} = 0 \qquad \ldots\ldots(3),$$

this being the sum of two open-pipe impedances in series. If either pipe be closed at one end, we substitute $-\cot$ for $+\tan$. The small discrepancies between theory and experiment are evidently due to the fact that end corrections are not included in the above formula.

In testing telephone lines it is often necessary to include the voice and ear in the system, particularly the latter since it forms an impedance closely coupled to the receiver. Artificial voices and ears such as those used by the Bell Telephone Co.[85] for testing may be constructed to simulate the physical characteristics of the organs, or the impedance of an average human ear may be measured by joining it on to an acoustic conduit. Figure 11 shows an apparatus devised for this purpose by Sivian and White[86]. In the diagram Z_d is the resistance of the ear-drum D, V_1 is the volume between it and a narrow tube whose volume is about 0·8 cm.³, V_2 is a volume, adjustable in magnitude by means of the sliding piece P, which can be put into communication with or be cut off from V_1 by sliding T over the holes H. If the combined impedance of V_1 and V_2 is Z_1, and Z_0 is the impedance of the source, a microphone C, measured at H looking towards C,

$$Z_0^{-1} = Z_d^{-1} + Z_1^{-1},$$

the impedances being in parallel. The pressure on D is proportional to the current I_0 in the microphone and to Z_0, which $= Z_d Z_1 (Z_D + Z_1)^{-1}$. Actually pairs of values of I_0 and Z_1 were obtained by making V_2 in succession 0, 10, and 40 cm.³ At 100 and 200 c./sec. and on the threshold of audibility the average values of Z_d, expressed as volumes of the equivalent air piston, worked out at 8·7 cm.³ and 2·7 cm.³ respectively. The anatomists give 1 cm.³ as the average actual volume of the space concerned.

Penman and Richardson[87] have used a purely acoustic method for determining

the impedance of a specimen of porous material, imitated by the more academic substitute of a bundle of narrow glass tubes, with the object of testing the Rayleigh theory of dissipation of sound in narrow tubes. Let the expression for the wave entering such a tube be written

$$\xi = \xi_0 \exp(-\alpha x) \exp\{i(\omega t - \beta x)\},$$

α being the absorption coefficient and β the phase-velocity coefficient for a pulsatance ω, equal to $2\pi n$. In the case of a tube of moderate width

$$\alpha = (rc)^{-1}(\omega\nu'/2)^{\frac{1}{2}} \quad \text{.....(4),}$$

$$\beta = \omega c^{-1} + (rc)^{-1}(\omega\nu'/2)^{\frac{1}{2}} \quad \text{.....(5),}$$

while for very narrow tubes

$$\alpha = \beta = \omega(rc)^{-1}(4\nu/\omega)^{\frac{1}{2}} \quad \text{.....(6).}$$

Here ν is the kinematic viscosity and ν' is the same quantity modified to allow for

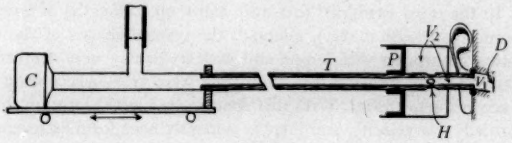


Figure 11.

heat-conduction. The criterion which separates the two cases is $r(i\omega/\nu)^{\frac{1}{2}}$, which for the first case should be less than 10, for the second case less than 1. The experimental tubes were border-line cases, and there was little difference between the results when the readings were substituted in either formula. It is on similar formulae that the Helmholtz-Kirchoff tube-correction to the velocity is based (compare § 3).

The quantity ρc (the product of the density and the velocity of sound in the medium, sometimes called the characteristic impedance) distinguishes the behaviour of sound at a boundary. Thus, if ξ_1 is the amplitude in a first medium with constants ρ and c and if the sound is incident *normally* on a second medium with constants ρ' and c', and if ξ_2 is the amplitude of the reflected wave, we have the relation, well known in connection with light, electro-magnetic waves and the like,

$$\xi_2/\xi_1 = (\rho'c' - \rho c)/(\rho'c' + \rho c) = (\rho'c'/\rho c - 1)/(\rho'c'/\rho c + 1) = (z-1)/(z+1), \text{ say.}$$

Thus, if we can measure the ratio of characteristic impedances z in air and in the porous substance, we can calculate the absorption coefficient of the specimen as $1-(\xi_2/\xi_1)^2$. In the first case we have, from equations (4) and (5),

$$z = 1 + (1-i)(\omega\nu'/2r^2)^{\frac{1}{2}} \quad \text{.....(7)}$$

and in the second, from equation (6),

$$z = (1-i)(4\nu/\omega r^2)^{\frac{1}{2}} \quad \text{.....(8).}$$

These are the relative impedances for infinite tubes, a correcting term being necessary for tubes of finite length.

The experiment consisted in driving a wide tube at one end from a loud-speaker, the bunch of narrow tubes being at the far end. Along the wide tube there will be maxima and minima of pressure-amplitude. It is easy to show that the impedance of the length l measured from one of these points, looking towards the unknown impedance, will match the value Z of the latter*. This impedance is, for a maximum, like that of the open pipe,

$$i\rho c \cot(\omega l_1/c),$$

and for a minimum

$$-i\rho c \tan(\omega l_2/c).$$

Hence

$$z = \cot(\omega l_1/c) \text{ or } -\tan(\omega l_2/c).$$

A determination of l_1 or l_2 will allow of a comparison with the theoretical values in equations (4) and (5), after separation of the real and imaginary parts of these expressions. The maxima and minima were found by means of a sliding section containing a manometric membrane, whose motions were amplified by a mirror, lamp and scale. In the cases examined (0·2-mm. tubes up to lengths of 6 cm. and frequencies from 250 to 900 c./sec.), although the general shapes of the theoretical curves relating absorption with length and with frequency were confirmed by experiment, the absolute values of the absorption were greater according to experiment than according to theory. With this abnormally high absorption there should go an abnormally low velocity, and there is evidently need for measurements of the velocity of sound in narrow tubes less than 1 mm. in diameter, much narrower than those used in the experiments described in § 3, before attempts to revise the theory can be made.

Kalyanaraman[88], using the tube method, gives results for porous materials indigenous to India which have not previously been determined. A number of physicists have been concerned recently with the relation between absorption of sound and the surface of non-porous materials. The following theoretical treatment is due to Paul Sabine[89]. Consider a rigid partition of mass m and area S separating two enclosures acted on by a forcing vibration $a \sin \omega t$. Writing the equation of motion in the familiar form

$$m\ddot{x} + \mu\dot{x} + kx = a \sin \omega t,$$

we obtain the solution for the response A of the forced vibrations of the partition in the form

$$A = [aS/m]\,[4\alpha^2 + (\omega_0^2/\omega^2 - 1)^2\,\omega^2]^{-\frac{1}{2}},$$

where α is the damping, equal to $\mu/2m$, and ω_0 is the natural pulsatance, equal to $(k/m)^{\frac{1}{2}}$, of the system. Regarding A as the amplitude of the partition sending waves into the second enclosure, the power of this source is proportional to $\omega^2 A^2$ and for the intensity I_2 on this side

$$I_2 \propto \omega^2 A^2/(\beta_2 c) \propto [a^2 S^2/(\beta_2 c m^2)]\,[4\alpha^2 + (\omega_0^2/\omega^2 - 1)^2\,\omega^2]^{-\frac{1}{2}} \quad \ldots\ldots(9),$$

β_2 being the total absorbing power of this enclosure. On the live side

$$I_1 \propto \omega^2 a^2 \propto a^2/(\beta_1 c), \quad \ldots\ldots(10),$$

* Incidentally, the pseudo-stationary waves set up in such a tube when it is terminated by an absorbent material present several interesting features, which have been examined by Irons[90], who has also designed a model for demonstration purposes.

β_1 being the total absorbing power on this side. Dividing relation (10) by relation (9) and assuming that acoustic conditions are kept the same in each enclosure (i.e. that $\beta_1 = \beta_2$) we get

$$I_1/I_2 \propto (m/S)^2 \, [4\alpha^2 + (\omega_0^2/\omega^2 - 1)^2 \omega^2]$$

and

$$\log I_1/I_2 = 2 \log (m/S) + \log [4\alpha^2 + (\omega_0^2/\omega^2 - 1)^2 \omega^2] + \text{constant} \ldots\ldots (11).$$

We see from equation (11) that, provided the natural frequency of the partition is small compared with that of the sound (the usual case), the logarithm of the reduction in intensity after transmission should bear a linear relation to the logarithm of the mass per unit area. This has been verified by experiments on a number of materials; figure 12 shows the results obtained by Davis[91, 92, 93] in this connection.

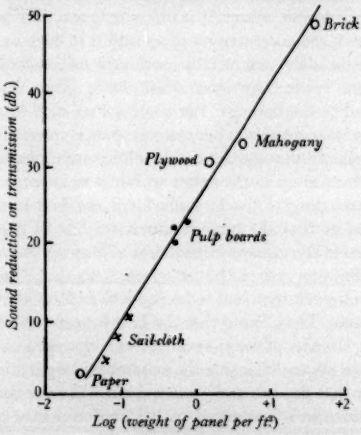


Figure 12.

The law has however its limitations. Where the difference between partitions is mainly one of weight, as between doors and masonry walls, it serves very well. But "walls of cinder concrete and slag tile show considerably greater reductions than do clay tile and plaster walls of equal weight" (Sabine).

Passing on to acoustic filters, we have a paper by Waetzmann and Noether[94] which describes experimental tests of the transmission shown by high-pass and low-pass filters for comparison with a modified theory. Whereas the original theory given by Stewart gave one definite cut-off frequency for each filter, the new theory indicates a series of maxima and minima stretching respectively above and below a certain frequency. There are in fact an infinite number of limiting frequencies. Only with an infinite number of equal sections will a filter completely cut off sound in a definite band of frequencies. In this research, the losses by transmission through a high-pass filter of six elements and a low-pass filter of four elements were deter-

mined, and the resulting (transmission, frequency) curves show the hills and valleys as predicted. However the authors are of the happy opinion that, in practice, a filter with a reasonable number of elements is covered by the simple Stewart theory with sufficient approximation. Two of Prof. Waetzmann's colleagues (Schuster and Kipnis[95]) have also examined exhaust-silencers in their aspect as filters. The ordinary baffle silencer consisting of a long tube divided off by baffles pierced with coaxal holes is essentially a low-pass filter, as Davis[96] points out. As the most considerable engine noises are coperiodic with the revolutions, or a small multiple thereof, it is not very efficacious. Let us consider under what conditions the silencer could be turned into a high-pass filter. In their earlier paper the Breslau scientists have shown that it is the filter constant $\beta = S_1/k_2 l_1$ which determines the attenuation in the rejected-frequency region. Here S_1 is the area, l_1 the length of each impedance in the main line, and k_2 the conductivity of an orifice in each branch. The attenuation is greater as β is smaller, and as it is more easy to construct a high-pass filter than a low-pass filter, with large attenuation for a given number of elements, attention was directed to the former. For small values of β the cut-off pulsatance $\omega = \omega_0 (1+4\beta)^{-1}$, where $\omega_0 = 2\pi c/2l_1$, i.e. the pulsatance corresponding to resonance ($\lambda = 2l_1$) in the element. Obviously, ω must lie above any pulsatances of the engine. Taking the highest frequency in the latter to be 2000 c./sec. and β equal to $\frac{1}{4}$, l_1 must be 4·25 cm. Reverting to the formula for β, we find that $S_1/k_2 = 1$, approx. It is suggested therefore that the total conductivity $\Sigma 2r$ of the orifices should be made practically equal to the cross-sectional area πR^2 of the main tube. Thus in one particular case the diameter $2R$ of the latter was 4·24 cm. The total conductivity of the orifices in each branch then had to be equal to $2\pi R$, or they could be replaced by a slit having this area. They found that the best form for the slit was a V with its vertex pointing along the axis of the silencer towards the open end. They also found that a number of filters in parallel inside the exhaust box were more potent than one. Incidentally, the fact that the exhaust gases are hot will raise the cut-off frequency by raising c, unless this rise is countered by the predominance of carbonic gases in the exhaust. Of course, one cannot attain silence without some sacrifice of efficiency in the engine. Since, as we know full well, certain engine-drivers, notably young motor-cyclists, will not make this sacrifice, the filter silencer will probably remain an ideal to be striven for.

Leaving these mundane applications of filters, we return to our academic groves to consider a theoretical paper on another type of filter which Stewart and his pupils once essayed, viz. a metal rod loaded at equal intervals with heavy metal collars, which acts as a low-pass filter to longitudinal waves passing through the rod. Lindsay and White[97] have examined this case and find that actually it is a band-pass filter, such that

$$\cos \omega l_1/c - [\omega m_2/(2\rho c S_1)] \sin \omega l_1/c$$

lies between ± 1 within the band. Here l_1 is the length of a main element, S_1 the area and ρ the density of the rod, and m_2 the mass of a collar. Thus the lowest cut-off frequency occurs when

$$\cot \omega l_1/c = \omega m_2/(2\rho c S_1),$$

above which there are alternate transmission and rejection bands. The string loaded at intervals and traversed by transverse waves, also the rod loaded with equal simple pendulums, are studied in detail. Such arrangements might be used as mechanical filters to prevent transmission of vibrations along shafts and the like.

§ 5. SUBJECTIVE ASPECTS OF SOUND

Dealing first with the voice we find that in a number of countries records of spoken language are being taken and analysed, mostly on string oscillographs. The principle is the same whether an actual Einthoven or Duddell oscillograph is used, or whether talking films in which light passes through a gate of variable width between two vibrating strips are made. Thus Scripture[98] and Lenk[99] have recorded English vowels, Obata[100, 101, 102] and pupils Japanese and Chinese speech sounds, Gemelli and Pastori[103] Italian vowels, and Backhaus[104] German vowels and consonants. The latter is more particularly interested in the starting noises which characterize both vocal and instrumental timbre. Scripture has produced these as isolated speech-bits by setting the mouth in the approximate position for a vowel with the lips closed over the thumb and suddenly withdrawing the latter. There seems to be general agreement with the *formant* theory of voice production. The mouth set in a certain position to produce a vowel brings out a certain range or formant of frequencies which are characteristic of that vowel, while the vocal cords are set to produce a buzzer note of a definite frequency. The resulting wave-form of the coupled system comprising the laryngeal cavities and the vocal cords shows the note of the latter as modulated by its passage through the mouth. In many cases, e.g. in the sound *ah*, the resulting wave shows a repeated pattern of diminishing peaks. Although Fourier analysis of the lower partials may be undertaken, it does not help us much in distinguishing the various speech sounds since the higher overtones are more characteristic of the formant than the lower, particularly in certain consonants. It is interesting to note that Miss Janvrin[105] has succeeded in diagnosing certain nervous affections in patients by the type of speech-record which they produce.

An interesting pathological experiment has been performed by Travis and Buchanan[106] in confirmation of the theory just outlined. They have passed a sinusoidal wave-form from a transmitter through a post-mortem preparation of the human head, and confirmed that higher overtones are added in transit. A number of writers have pointed out that damping must be introduced by the soft material with which the linings of the throat, etc. are coated, and that this damping may not be the same for all overtones.

Most of the work in the science of hearing of late has been concerned with the physical measurement of that complex physiological and psychological sensation known as *loudness*. As we should expect, there is considerable variation between the results of different workers in this sphere. Some have failed to take proper precautions to ensure that the subject's attention shall not be diverted by extraneous sights and sounds. The physicist is usually fond of his gadgets, and does not consider that

their operation may distract the subject. Others have not inflicted their tests on a sufficient number of patients to ensure that their results shall represent what they call a normal ear.

Dealing first with the measurement of minimum audibility, we must notice some careful measurements by Huizing*. A transmitter connected to a valve oscillator giving simple-harmonic motion had its amplitude measured in two ways, directly by means of a mirror attached to the diaphragm, and by making the vibrating plate form part of a condenser microphone. In this latter way amplitudes down to 1 $\mu\mu$ (10^{-7} cm.) could be detected. After investigation of the sound-field in the neighbourhood of the source, the amplitude corresponding to the minimum audibility for ears at 20 cm. and various frequencies was determined. The ear was most receptive at a frequency of 800 c./sec., when the calculated energy-flow on to the ear was 10^{-9} erg/cm^2-sec. Intensity-values of one-tenth of this quantity could be measured with accuracy. As far as one can tell, with a very scanty knowledge of the Dutch language, the effect of the subject's head on the sound-field is ignored in the calculation of threshold loudness. Following on a detailed examination by Geffcken and Keibs[107] of the thermophone (an instrument due originally to Wente in which an alternating current passing through a metal strip or wire produces weak sound waves from the expansions and contractions accompanying the periodic heating and cooling) the first author has measured with Waetzmann[108] this same elusive quantity. The instrument possesses the merit of a reliable and stable calibration. Instead of a steady fall of threshold loudness from the lower pitch-limit to 1000 c./sec. which earlier observers give, the new research indicates a fall from 100 to 800 (cf. Huizing) followed by three peaks and troughs centering round the same average. Whereas Huizing gives rather higher values than other workers, Waetzmann and Geffcken find decidedly lower values, which, considering that both researches were obviously carried out with all possible care, bears out what was said above with regard to the uncertainties accompanying such measurements.

A number of workers have been busy with the validity of the Weber-Fechner law (according to which sensation is proportional to the logarithm of the stimulus) in connection with the change of loudness L which the mind associates with a given change in physical intensity E. If $L = k \log E$, and E be increased to E^m,

$$mL = k \log (E^{m+1}/E),$$

or

$$\exp \{(m-1)\, L/k\} = E^m/E.$$

If then we raise or reduce the intensity in a certain ratio, i.e. by so many decibels, a bel being a ratio of intensities, and if we let a listener estimate the percentage increase or decrease in loudness, on plotting one against the other as in figure 13 (from Geiger and Firestone[109]) we should get an exponential curve. The three curves shown are for three different original levels 30, 55, and 80 db. above the threshold, and though of the right shape indicate that k depends on the original value of E. Ham and Parkinson[110] obtain similar data using warble tones instead of a constant frequency, while Riesz[111, 112, 113, 114, 115] and others have investigated the matter

* Thesis for doctorate (J. B. Wolters, Gröningen, 1932).

from the point of view of the least perceptible increase of intensity. His method was to determine the number of steps δE which the listener could distinguish as the tone was raised step by step from intensity E_1 to E_2. The loudness-sensitivity is then given by $\delta E/E$. The number of steps is shown on a graph against decibel level above threshold for a number of frequencies. The number of steps increases with level in a roughly exponential manner. If now we select, from these curves for various levels, that level at which the number of distinguishable steps is reduced by a half, and graph the two levels, we get a straight line for all frequencies, showing that the reduction of loudness is proportional to the original intensity of the sound.

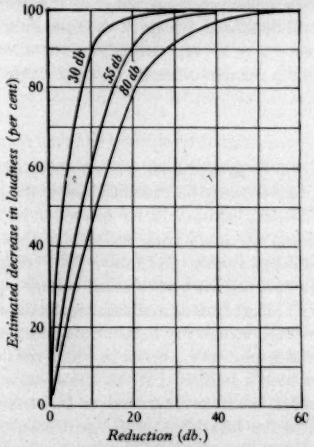


Figure 13.

In fact, with adequate control the Weber-Fechner law in acoustics is vindicated. For a more detailed account of the whole question of loudness in theory and experiment the reader is referred to a paper by Fletcher and Munson[116].

Von Bekesy[117, 118] has investigated the loudness of sounds lasting a very short time (down to $\frac{1}{8}$ sec.). He opines that since in listening we are usually concerned with impulses such as the starting noises to which Backhaus has drawn attention rather than with long-continued sounds, it is important to obtain data on this aspect of hearing, for application to acoustics of buildings. The same author has investigated the behaviour of the ear itself under such conditions. Subjects were trained to give their ear-drums and associated ossicles an impulse by sudden expulsion into the middle ear of compressed air from the Eustachian tube, an anatomical trick which he assures us is easily acquired though painful at first. Our long-suffering subject has his ear-drum connected through a tube 1·5 metres long to a condenser microphone and recording device which is cut out before reflected waves in the tube

can reach it. From the oscillograph records of the damped vibrations of the ear-drum it is readily possible to calculate the damping-coefficient and natural frequency of the middle ear. The latter is about 1200 c./sec., i.e. about the same frequency for which the ear shows maximum loudness-sensitivity. Tones lasting a short time and simultaneously impressed on the two ears gave the sensation of binaural location even when their frequencies differed considerably.

Another binaural-location research, that of Reich and Behrens[119], is concerned with a tone-mixture or complex note. They found that the direction-sensitivity depended on the relative rate of change of pressure at the two ears. As long as there were corresponding peaks in the pressure-change at the two ears the direction of a complex source could still be sensed, provided the time interval between the arrival of the peaks did not exceed 6×10^{-4} sec. In judging the direction of a noise the ear picks out and estimates the relative intensity of those components which have the smallest time-differences.

§ 6. TECHNOLOGY

Under this heading will be grouped some researches, having industrial applications, which could not conveniently be classified under the preceding headings.

The testing of reproducing-equipment for distortion is now an important process in the gramophone and wireless trades, and special apparatus has been devised for carrying out this test expeditiously. Ballantine[120, 121] employs the following apparatus. Current from a beat oscillator at constant voltage but variable frequency is fed to the loud-speaker. Voltage from a condenser microphone on which the sound falls is led after suitable amplification to a logarithmic voltmeter. The basic element of this voltmeter is a variable-mu valve, i.e. one in which the pitch of the control grid varies along the grid in such a fashion that the anode-current, grid-bias curve no longer follows the $\frac{3}{2}$-power law of ordinary valves but a law in which the amplification is proportional to the logarithm of the grid-bias. After rectification and smoothing the resulting d.c. output is impressed upon the control grids and automatically keeps the output voltage of the amplifier constant. A voltmeter connected across the grid will then indicate the logarithm of the input voltage. The condenser which varies the frequency of the oscillator is shaped in the usual fashion for securing that the logarithm of the frequency is proportional to the rotation, accomplished by a small motor which also rotates a drum carrying a sensitized paper strip. By means of a beam of light and a mirror the voltmeter recordings are traced in an axial direction on the drum. Thus the apparatus automatically traces the (response, frequency) curve in log-log form.

In the apparatus used by Davis[122] the voltage from the microphone is taken after amplification to one pair of electrodes of a cathode-ray oscillograph. To the other pair of plates at right-angles to the first is led from the valve oscillator a d.c. voltage whose value is proportional to the logarithm of the frequency. This is accomplished by passing the output through a frequency-weighting network and amplifier.

Thus the electron spot in the viewing screen receives a horizontal displacement proportional to the frequency and a vertical one proportional to the output of the

microphone. As the oscillator is taken through the gamut, the spot on the screen traces out the response curve (which can be photographed) of the loud-speaker as registered by the distortion-free microphone. The frequency-weighting network to which reference has been made consists essentially of a resistance-capacity stage of amplification in which the coupling condenser is small. Consequently the amplification factor increases nearly in proportion to the frequency. In the rectifier which follows, the d.c. output is made nearly proportional to the logarithm of the frequency.

A number of new types of reproducer have appeared recently. There is first the piezoelectric transmitter, in which, as its name indicates, the diaphragm is driven by quartz oscillators. It is suggested that one of these should be used in tandem with a moving-coil loud-speaker to yield a better response in the high-frequency range where the response of the loud-speaker falls off. Gemant[123,124,125] has succeeded in modulating the d.c. corona discharge (electric wind) by sound-waves in his *Glimmlautsprecher*.

We must also mention here the process sometimes known as the Blattnerphone, whereby speech or music may be recorded by magnetic means on a steel strip and reproduced later. This is now frequently used by the British Broadcasting Corporation when a temporary record of a speech is required during broadcasting, to be radiated later in the day in another programme. Although this official adoption is recent, the process dates back to the early days of wireless telegraphy and was a device used for tape recording by Poulson. Erwin Meyer and Schüller[126] have investigated this apparatus from the point of view of its performance. The continuous steel band is carried round on pulleys in front of a coil through which an alternating current engendered by the speech or music passes. This magnetizes the steel in the direction of its motion in a corresponding fashion. When reproduction is desired, the band is passed in front of another coil with permalloy core and the induced current is fed into an amplifier. The record can be erased, so that the band is ready for subsequent use, by passing the band near a sufficiently strong direct or alternating magnetic field so that the first saturates the iron and the latter demagnetizes it by taking it through hysteresis cycles of successively smaller amplitudes as the band retreats from its influence. In the research to which we are referring, frequency-response curves were taken and the falling off at high frequencies was discussed. This latter effect is analogous to the difficulties in talking films when one attempts to record wave-lengths comparable with the slit-widths, because interference phenomena intervene. Amplitude-distortion occurs owing to the hysteresis in the steel band. Listeners to a Blattner broadcast will have noticed these defects. Another defect is the considerable amount of ground noise, due to inequalities in the band surface and mostly of low frequency, 100 to 1000 c./sec. Writing of broadcasting reminds us of the electrical music which is now a feature of Continental relays. We have already mentioned the pianoforte with electrical pick-up. Most of the instruments which claim to produce music from the aether directly are, however, valve oscillators giving only one note at a time in which the capacity of one of the condensers is varied to alter the frequency either by pressing keys or, more simply,

by approaching one hand while the other hand works a volume control. The quality is ultimately limited by the loud-speaker to which the electrical oscillations are led, and as those without keyboard, which are more common and cheaper, exhibit a distressing whining as the player literally feels his way from one note to the next, they are scarcely likely in their present stage of development to appeal to the serious musician.

We should also add to this list the unidirectional ribbon microphone of Weinberger, Olson and Massa[127]. They take the velocity or pressure-gradient microphone, a light corrugated metallic ribbon suspended in a magnetic field and open to the air on both sides, in which a difference of pressure on the two sides gives rise to an e.m.f. proportional to the particle-velocity in the sound wave. To this is added an ordinary pressure-operated microphone. The relation between intensity-distribution and inclination to the normal is given by $I = I_0 \cos\theta$ for the first and $I = I_0$, a constant, for the second. Combination of the two principles should give us a cardioid distribution $I = I_0(1 + \cos\theta)$ which means a wide pick-up angle all in front instead of one having the bidirectional characteristics that the first type has. Actually, as the theory shows, the two principles can be combined in one instrument in which one ribbon is employed, half of it being velocity-operated and the other half pressure-operated. In practice this means that the upper half is open at back and front, but that the lower half is backed by a pipe connected to a damping labyrinth in a box below the instrument. The object, in theory, of this rather strange device is to simulate a pure resistance load at all frequencies. A long heavily lagged pipe would be the best approach to this type of load, but is replaced by the described device for obvious reasons. Such a microphone should be very useful on the theatre or opera stage where there is a desire to broadcast the performance and exclude the concurrent chatter from the stalls!

We pass on to the subject of building acoustics. Here again, although the original subjective method is still used, the current tendency is to replace the ear and stop-watch by instruments which automatically delineate the decay of reverberation in a room. Norris[128] describes improvements in the apparatus designed by himself and Andree for measuring the absorption of a specimen from a record of the reverberation in its vicinity. Source, recorder and panel are placed in a small chamber 5 ft. × 6 ft. × 9 ft. The microphone and loud-speaker are located at the centre of a paddle wheel which rotates and mixes the sound, as the engineer says; that is to say, it keeps the interference patterns stirred up. The paddle is large, nearly filling the chamber. In the recording-apparatus, which is of course outside the chamber, a cam-operated device turns on the source, shuts off the source and connects the oscillograph to the microphone at any set interval of time later. This series of operations is carried on continuously with a period of 3·6 sec. By shifting the phase cam the decay curve can be plotted step by step. Interference phenomena in reverberation are considered by Barrow[129], who prefers a source of oscillating frequency, or warble tone, as also does Hunt[130, 131], while von Bekesy[132] is concerned with the absorption which is of necessity produced in passing along the surface of an absorbent, such as is formed for a listener in a theatre by the audience in front of him.

The other chief progress in reverberation is the recognition that it consists essentially of the free damped vibration of the air in the room. Knudsen[133], among others, has thrown light on this point. The natural frequencies of a room measuring $l_1 \times l_2 \times l_3$ are found by giving p, q and r integral values (or zero) in the formula

$$n = (c/2)\,(p^2/l_1^2 + q^2/l_2^2 + r^2/l_3^2)^{\frac{1}{2}}.$$

These frequencies can of course be picked up by an oscillator of variable tuning and a microphone in the room in question. The graver modes (found by giving p, q, r, in turn the value 1, while the others are 0) are readily recognized as waxings and wanings in the oscillogram of the reverberation, especially if a pair are sufficiently near together to produce a slow beat. In one room examined by Knudsen a beat of 7·0 c./sec. between two frequencies of 92·8 and 99·8 c./sec. was clearly discernible in the decay curve, no matter what source frequency was used. Other technical papers on reverberation too numerous to mention here will be found in the *Journal of the Acoustical Society of America*.

In transmission, besides those which we have mentioned under the heading of impedance, there are papers by Ritchie[134] and by Wintergerst and Knecht[135] on the amount of sound transmitted by apertures, such as open windows, of different sizes. The former writer constructed a partition separating a chamber into two parts, the one containing an organ pipe swinging to and fro in pendulum fashion as source. Behind an aperture in the partition was the microphone in a non-reflecting enclosure. When the aperture can be considered as a circular hole in a thin plate the transmission should be proportional to the diameter. This was verified for apertures, and a modified theory, due to Lamb, which came under review—viz. that the transmission through a circular hole is $8/\pi^2$ times the energy-flux through an equal area of the primary waves—was confirmed by measurement. The same author has obtained data for several apertures in parallel in the same partition and for others formed of short tubes. It was found in the latter case that the transmission suffered an exponential decrease with increase of length of tube. One ought to mention, however, that such a large-scale method of determining the conductivity of orifices and necks raises difficulties in that the stationary wave pattern in the primary chamber may be altered with change in aperture, and that Mr Ritchie found evidence of interference in spite of the swinging of the source, in that the indications of the microphone depended largely on its position in the receiving chamber. Wintergerst and Knecht made their partition in a long wide pipe, but otherwise the methods were similar. Openings 1 to 10 mm. in diameter were used; also narrow tubes up to 40 cm. long, the latter showing stationary waves at the appropriate frequencies.

In this connection, Gemant[136,137] has measured the hydrodynamic resistance of some specimens of porous acoustic material to steady one-way flow. Although this resistance is no doubt related to the transmission coefficient for sound, yet there is such a difference between the mechanism of steady and alternating resistance that to the writer it seems scarcely profitable to attempt to correlate them.

Apparatus for the measurement of noise, that is, according to the Acoustical

Society's definition, "any unwanted sound," has now become a standard equipment for engineering firms, and several commercial noise-meters are now on the market, although the simple method devised by Davis of listening to the decaying sound of an ordinary tuning fork until it is masked by the noise in question still continues in effective use[138]. In an apparatus made up by the General Radio Co., the noise picked up by a carbon microphone is amplified in such a fashion that most weight is given to frequency components in the middle of the scale, to which the ear is most sensitive. A vacuum-tube voltmeter or an oxide rectifier gives the overall electric power output from the amplifier, which, if the meter is calibrated in some way, is a measure of the noise. The standard usually adopted for calibration is the threshold of audibility for a normal ear of a tone of 1000 c./sec., obtained from a valve oscillator and loud-speaker. The latter is placed at the same distance from the microphone of the noise-meter as is the listener's ear, and the volume of its output is throttled down until it becomes inaudible. The corresponding reading of the meter is then the standard, and other noises are measured in decibels by the ratio of their overall intensities to this. Another type, such as that of the Western Electric Co., uses the masking effect. Difficulties arise in both types when one seeks to compare noises of differing average-pitch range. It will obviously be difficult to get adequate comparison between, say, the low-pitched rumble of the passage of an underground train and the high-pitched noise of an aeroplane on the basis of a single pure tone of fixed frequency. To get over this difficulty, the noise is masked by one of a number of warble tones in turn, whose intensity is raised until it is just audible above the noise. The frequency limits of the eight warbles in the instrument under review are 100, 200, 400, 800, 1600, 3200, 6400, 10,000, so that they cover between them the greater part of the musical scale. They are produced by a beat-tone oscillator, in which one of the small condensers has its movable plates continuously rotated four times a second. The noise is measured by the electrical input to a loud-speaker, equipped with volume control, at the inception of masking, for each warble in turn. The reader will find a detailed criticism of the principal instruments on the market and a review of the possible errors arising in their use in papers by Baron[139] and by Obata and Morita[140]. M. Baron prefers the masking type with warble tone because of the difficulty we have mentioned above of comparing a noise and a pure tone, severally excited. And if, as he says, instead of a single reference tone one employed a number of different tones to get over this difficulty, systematic errors would mount up and *on serait tout à fait dérouté*.

Osbon and Oplinger[141] have devised a useful portable meter in which the frequency-range can also be analysed. This is done by superposing on the output from the microphone an electrical oscillation of fixed frequency between 6800 and 12,000 c./sec. with which it produces combination tones. When one of these combination tones is in resonance with that (6800) of a brass rod, supported at its centre, it is picked up by a coil and passed to the output meter. Retuning the oscillator will produce the same combination tone from another primary in the noise and so on.

It is of course not easy to analyse a noise in a profitable way—by which one means, so as to obtain some indication of the origin of the frequency components—

except in rotary machines. At present the noise of aeroplanes seems to be attracting the attention of those whose work lies in technical acoustics, and as our allotted space is running out it is only researches in this direction which can gain our attention here*.

Researches have been made by Paris[147, 148] and Kemp[149] under the aegis of the Aeronautical Research Committee into the distribution round an aeroplane, both on the ground and in flight, of the low-frequency noise from airscrews. It is well known that the fundamental airscrew noise is given by the product of the number of blades and the revolutions per second. In his first paper Paris modifies the theory of Lyman and Webb, who dealt with the problem as that of the substitution for the origin of noise of a number of sources and sinks placed symmetrically round the airscrew disc. Since the experimental curve is not symmetrical about the plane of rotation, such a distribution is too simple. Paris finds that a ring of sources at a specified distance in front of this plane, another in the plane, and a ring of sinks at the same distance behind gives the representation required. The distinction between source and sink is merely one of phase—the sink lags behind the corresponding source by half a period. The velocity-potential of these gives, like the experiments, an oval distribution curve for the sound, with a maximum (along the major axis of the oval) inclined 20° forward of the plane of rotation. The experimental work was done with a single-engined bomber having a twin-bladed airscrew rotated at 1750 rev./min., and a hot-wire microphone tuned to 27 c./sec. In the second paper experiments on the second harmonic of the exhaust of a twelve-cylinder **V** engine are recorded, and these show that most of the sound is projected forward, while the amount sent back along the slipstream is relatively small. Finally, as it was thought that the results might be affected by local reflections when the aeroplane was on the ground, a pilot flew it overhead, and the sound (this time, again, the engine fundamental) was picked up by a tuned hot-wire microphone on the ground, the aeroplane being spotted meanwhile in a camera obscura. These results confirmed the earlier ones as to the direction in which the fundamental was radiated. In Kemp's set-up, a condenser microphone was used; the fuselage was fixed vertically and the exhaust noise was baffled by being passed into a pit filled with steel shavings and covered with earth. The various harmonics were picked out by resonance methods before the output from the amplifier was measured. The distribution curves of Paris were confirmed; in addition, as we should expect, it was found that the amplitude in the first six harmonics fell off rapidly.

At the same time Obata, Morita and Yosida have been taking sound-records of full-sized and of model airscrews at the Tokyo Aeronautical Research Institute[150, 151, 152]. They take complete oscillograph records (that is, without resonators) and attempt to apply Fourier analysis to these, a difficult task since, owing probably to casual fluctuations in the speed of the engine, the records do not exhibit a periodicity exactly recurring. Further, the model airscrews were mounted near the laboratory

* The theory of silencers has been dealt with in § 4. Other papers dealing with special types of noise within the period are given in the References as follows: lorries[142], ships[143], bridges[144], New York L[145], gears[146].

wall and therefore the noise was liable to interference through reflections. Their method has however the merit of including the airscrew frequencies without the exhaust frequencies, the models being rotated electrically. Thus were included both the blade sounds and the higher pitched aeolian tones formed as the blades cleave the air, which give rise to the swishing noise associated with airscrews. From the point of view of annoyance, the latter noises are quite as important as, if not more important than, the exhaust and blade noises, since they abut on the pitch region of maximum sensitivity of the ear.

Echo-sounding continues to flourish as a branch of applied acoustics. Messrs Hughes now have an instrument of the Langevin type in which magnetostriction replaces piezoelectricity as the source of a train of waves sent down to the sea-bed and picked up on return. Hubbard[153] and Dorsey[154] consider how sound-ranging may be used for finding direction under water. Engineers are now turning their attention to echo-sounding from the air. Eisner and Kruger[155] have tested an echolot which emits and detects an impulse from a cartridge fired overboard, with a view to giving the pilot information of the type of terrain below him in fog, as well as his height. They take the reflection coefficient of water as 1, of ice as 1·07, of meadow land as 0·49, and of woods as 0·21 to 0·45. A group of houses gives multiple echoes. It seems, however, rather doubtful whether a measurement of the reflection coefficient will be of much use to the pilot when he is in the air, even if it could be made with sufficient accuracy. Florisson[156] has tried a jet of compressed air at the vertex of a cone as source. The pilot is apparently expected to listen through a tube for the echo, and himself to work a chronometer graduated in metres of height. Again we should be rather doubtful whether the pilot could readily manipulate such an apparatus amid the roar of the engine. M. Florisson has worked the apparatus himself as passenger, at heights varying from 10 to 240 metres.

§ 7. ACKNOWLEDGMENTS

The writer wishes to thank Profs. Andrade, Biquard, and Edgar Meyer for the loan of the photographs reproduced.

REFERENCES

(1) Kneser. *Ann. Phys.*, Lpz., **11**, 761, 777 (1931).
(2) Groszmann. *Phys. Z.* **33**, 202 (1932).
(3) —— *Ann. Phys.*, Lpz., **13**, 681 (1932).
(4) Henry. *Proc. Camb. phil. Soc.* **28**, 249 (1932).
(5) Kneser and Zühlke. *Z. Phys.* **77**, 649 (1932).
(6) Richards and Reid. *J. chem. Phys.* **1**, 114, 737 (1933).
(7) Teeter. *J. chem. Phys.* **1**, 251 (1933).
(8) Richards. *J. chem. Phys.* **1**, 863 (1933).
(9) Kneser. *Ann. Phys.*, Lpz., **16**, 337 (1933).
(10) Eucken and Becker. *Z. phys. Chem.* B, **20**, 467 (1933).
(11) Barnes. *J. acoust. Soc. Amer.* **3**, 579 (1932).
(12) Luck. *Phys. Rev.* **40**, 440 (1932).

(13) GRABAU. *J. acoust. Soc. Amer.* **5**, 1 (1933).
(14) PIELEMEIER. *Phys. Rev.* **41**, 833 (1932).
(15) RICHARDSON. *J. acoust. Soc. Amer.* (abstract), **4**, 177 (1933).
(16) BÜCKS and MÜLLER. *Z. Phys.* **84**, 75 (1933).
(17) HUBBARD. *J. acoust. Soc. Amer.* **4**, 172 (1932).
(18) HERSHBERGER. *J. acoust. Soc. Amer.* **4**, 273 (1933).
(19) BOURGIN. *J. acoust. Soc. Amer.* **4**, 108 (1932).
(20) KNUDSEN. *J. acoust. Soc. Amer.* **5**, 122 (1933).
(21) KINOSHITA and ISHTI. *Rep. Inst. phys. Chem. Tokyo*, **19**, 83 (1932).
(22) SWANSON. *Rev. sci. Instrum.* **4**, 603 (1933).
(23) BLANCANI and DOGNAN. *C.R. Acad. Sci.*, Paris, **197**, 1070 (1933).
(24) MALOV. *Hochfrequenztech. u. Elektroakust.* **42**, 115 (1933).
(25) DEBYE and SEARS. *Proc. nat. Acad. Sci.*, Wash., **18**, 409 (1932).
(26) LUCAS and BIQUARD. *C.R. Acad. Sci.*, Paris, **194**, 2132 (1932).
(27) —— —— *C.R. Acad. Sci.*, Paris, **195**, 121 (1932).
(28) —— —— *J. Phys. Radium*, **3**, 464 (1932).
(29) BÄR and EDGAR MEYER. *Phys. Z.* **34**, 393 (1933).
(30) HIEDEMANN and ASBACH. *Phys. Z.* **34**, 494, 734 (1933).
(31) BERGMANN. *Phys. Z.* **34**, 761 (1933).
(32) BIQUARD. *C.R. Acad. Sci.*, Paris, **196**, 257 (1933).
(33) —— *C.R. Acad. Sci.*, Paris, **197**, 309 (1933).
(34) DYE. *Proc. roy. Soc.* **138**, 1 (1932).
(35) OSTERBERG. *Phys. Rev.* **43**, 819 (1933).
(36) STRAUBEL. *Phys. Z.* **34**, 894 (1933).
(37) RÖHRICH. *Z. Phys.* **73**, 813 (1932).
(38) GIEBE and BLECHSCHMIDT. *Ann. Phys.*, Lpz., **18**, 417 (1933).
(39) KRÖNCKE. *Z. tech. Phys.* **13**, 196 (1932).
(40) SZALAY. *Z. phys. Chem.* A, **164**, 234 (1932).
(41) MAHAJAN. *Indian J. Phys.* **7**, 539 (1933).
(42) SAWADA. *Z. tech. Phys.* **14**, 353 (1933).
(43) YAMASHITA and AOKI. *Mem. Coll. Sci. Kyoto*, **16**, 367 (1933).
(44) DERJAGUIN. *Phys. Z. Sowjet.* **3**, 574 (1933).
(45) OLIVER. *Phil. Mag.* **14**, 318 (1932).
(46) ZAHRADNICEK and ZAK. *Ann. Phys.*, Lpz., **12**, 662 (1932).
(47) COLWELL. *J. Franklin Inst.* (several papers).
(48) MCLACHLAN. *Proc. phys. Soc.* and *Phil. Mag.* (several papers).
(49) FLÜGGE. *Z. tech. Phys.* **13**, 199 (1932).
(50) SEIFFERT. *Z. Phys.* **76**, 407 (1932).
(51) VON BEKESY. *Z. Phys.* **79**, 668 (1932).
(52) HALL. *J. acoust. Soc. Amer.* **4**, 83 (1932).
(53) O'NEIL. *J. acoust. Soc. Amer.* **3**, 483 (1932).
(54) SACERDOTE. *Alta Frequenza*, **2**, 516 (1933).
(55) MARY WALLER. *Proc. phys. Soc.* **45**, 101 (1933).
(56) JONES and ALDERMAN. *J. acoust. Soc. Amer.* **4**, 331 (1933).
(57) CURTISS and GIANNINI. *J. acoust. Soc. Amer.* **4**, 245 (1933).
(58) —— *J. acoust. Soc. Amer.* **5**, 159 (1933).
(59) KRÖNCKE. *Phys. Z.* **33**, 733 (1932).
(60) MINNAERT. *Phil. Mag.* **16**, 235 (1933).
(61) NIPPOLDT. *Z. Geophys.* **8**, 403 (1932).
(62) ABBÉ CARRIÈRE. *Rev. Acoust.* **2**, 333 (1933).
(63) VOLKOVSKY. *C.R. Acad. Sci.*, Paris, **197**, 902 (1933).
(64) THOM. *Proc. roy. Soc.* **141**, 651 (1933).
(65) MARTY. *J. Phys.*, Paris, **4**, 551 (1933).
(66) TYLER. *Phil. Mag.* **16**, 504 (1933).
(67) BROWN. *Phil. Mag.* **13**, 161 (1932).
(68) ZICKENDRAHT. *Helv. phys. Acta*, **5**, 317 (1932).

(69) Andrade. *Proc. roy. Soc.* **134**, 445 (1931).
(70) —— *Philos. Trans.* **230**, 413 (1932).
(71) Schlichting. *Phys. Z.* **33**, 327 (1932).
(72) Abbé Carrière. *J. Phys. Radium*, **3**, 355 (1932).
(73) Rilbentrop. *Z. tech. Phys.* **13**, 356 (1932).
(74) Auger. *C.R. Acad. Sci.*, Paris, **195**, 516 (1932).
(75) Obata and Tesima. *Proc. phys.-math. Soc. Japan*, **15**, 125 (1933).
(76) Sato. *Rep. aero. Res. Inst. Tokyo*, **93**, 339 (1933).
(77) —— *Proc. phys.-math. Soc. Japan*, **14**, 363 (1932).
(78) Kaye and Sherratt. *Proc. roy. Soc.* **141**, 123 (1933).
(79) Vance. *Phys. Rev.* **39**, 737 (1932).
(80) Van Itterbeek. *Rev. Acoust.* **2**, 81 (1933).
(81) Canac. *Rev. Acoust.* **1**, 52 (1932).
(82) Rocard. *C.R. Acad. Sci.*, Paris, **195**, 112 (1932).
(83) Flanders. *Bell Syst. tech. J.* **11**, 402 (1932).
(84) Bate. *Phil. Mag.* **16**, 562 (1933).
(85) Bell Telephone Co. *Bell Syst. tech. J.* **11**, 293 (1932).
(86) Sivian and White. *J. acoust. Soc. Amer.* **4**, 288 (1933).
(87) Penman and Richardson. *J. acoust. Soc. Amer.* **4**, 322 (1933).
(88) Kalyanaraman. *Indian J. Phys.* **7**, 229 (1932).
(89) Paul Sabine. *J. acoust. Soc. Amer.* **4**, 38 (1932).
(90) Irons. *Rev. Acoust.* **2**, 211 (1933).
(91) Davis. *Phil. Mag.* **15**, 309 (1933).
(92) Bedell and Swartzel. *J. acoust. Soc. Amer.* **5**, 34 (1933).
(93) Wintergerst and Klupp. *Z. Ver. dtsch. Ing.* **77**, 91 (1933).
(94) Waetzmann and Noether. *Ann. Phys.*, Lpz., **13**, 212 (1932).
(95) Schuster and Kipnis. *Ann. Phys.*, Lpz., **14**, 123 (1932).
(96) Davis. *Phil. Mag.* **16**, 787 (1933).
(97) Lindsay and White. *J. acoust. Soc. Amer.* **4**, 155 (1932).
(98) Scripture. *J. acoust. Soc. Amer.* **5**, 148 (1933).
(99) Lenk. *S.B. Akad. Wiss. Wien*, 2A, **142**, 313 (1933).
(100) Obata and Satta. *Proc. phys.-math. Soc. Japan*, **14**, 341 (1932).
(101) Obata and Tesima. *Jap. J. Phys.* **8**, 1 (1932).
(102) —— —— *Jap. J. Phys.* **9**, 510 (1933).
(103) Gemelli and Pastori. *Rev. Acoust.* **2**, 169 (1933).
(104) Backhaus. *Z. tech. Phys.* **13**, 31 (1932).
(105) Janvrin. *Lancet*, p. 1384 (1932).
(106) Travis and Buchanan. *Science*, **77**, 121 (1933).
(107) Geffcken and Keibs. *Ann. Phys.*, Lpz., **16**, 404 (1933).
(108) Waetzmann. *Phys. Z.* **34**, 234 (1933).
(109) Geiger and Firestone. *J. acoust. Soc. Amer.* **5**, 25 (1933).
(110) Ham and Parkinson. *J. acoust. Soc. Amer.* **3**, 511 (1932).
(111) Riesz. *J. acoust. Soc. Amer.* **4**, 211 (1933).
(112) Marvin. *J. acoust. Soc. Amer.* **3**, 388 (1932).
(113) Laird, Taylor and White. *J. acoust. Soc. Amer.* **3**, 393 (1932).
(114) Lifshitz. *J. acoust. Soc. Amer.* **5**, 31 (1933).
(115) Knauss and Hale. *Rev. sci. Instrum.* **4**, 447 (1933).
(116) Fletcher and Munson. *J. acoust. Soc. Amer.* **5**, 82 (1933).
(117) Von Bekesy. *Ann. Phys.*, Lpz., **16**, 844 (1933).
(118) —— *Phys. Z.* **34**, 577 (1933).
(119) Reich and Behrens. *Z. tech. Phys.* **14**, 1 (1933).
(120) Ballantine. *J. acoust. Soc. Amer.* **5**, 10 (1933) (abst. **4**, 10 (1932)).
(121) Fay and Hall. *J. acoust. Soc. Amer.* **5**, 46 (1933).
(122) Davis. *Phil. Mag.* **15**, 309 (1933).
(123) Gemant. *Z. Phys.* **78**, 240 (1932).
(124) McLachlan. *J. acoust. Soc. Amer.* **5**, 107 (1933) (electrostatic loud-speaker).

(125) BINDER. *Phys. Z.* **33**, 85 (1932) (electrodynamic loud-speaker).
(126) ERWIN MEYER and SCHÜLLER. *Z. tech. Phys.* **12**, 75 (1932).
(127) WEINBERGER, OLSON and MASSA. *J. acoust. Soc. Amer.* **5**, 139 (1933).
(128) NORRIS. *J. acoust. Soc. Amer.* **3**, 361 (1932).
(129) BARROW. *J. acoust. Soc. Amer.* **3**, 562 (1932).
(130) HUNT. *J. acoust. Soc. Amer.* **5**, 127 (1933).
(131) HOPPER. *J. acoust. Soc. Amer.* **3**, 415 (1932).
(132) VON BEKESY. *Z. tech. Phys.* **14**, 6 (1933).
(133) KNUDSEN. *J. acoust. Soc. Amer.* **4**, 20 (1932).
(134) RITCHIE. *J. acoust. Soc. Amer.* **3**, 402 (1932).
(135) WINTERGERST and KNECHT. *Z. Ver. dtsch. Ing.* **76**, 777 (1932).
(136) GEMANT. *S.B. preuss. Akad. Wiss.* **27**, 579 (1933).
(137) KÜHL and MEYER. *S.B. preuss. Akad. Wiss.* **26**, 406 (1932).
(138) STOWELL. *J. acoust. Soc. Amer.* **4**, 344 (1933).
(139) BARON. *Rev. Acoust.* **2**, 116, 189 (1933).
(140) OBATA and MORITA. *J. acoust. Soc. Amer.* **4**, 129 (1932).
(141) OSBON and OPLINGER. *J. acoust. Soc. Amer.* **5**, 39 (1933).
(142) MEYER and WILLMS. *Z. Ver. dtsch. Ing.* **76**, 1 (1932).
(143) RICHARDSON. *Trans. N.-E. Cst Instn Engrs Shipb.*
(144) VASSILIÈRE-ARLHAC. *Rev. Acoust.* **2**, 225 (1933).
(145) STANTON and TWEEDALE. *J. acoust. Soc. Amer.* **3**, 371 (1932).
(146) ABBOTT. *J. acoust. Soc. Amer.* **3**, 445 (1932).
(147) PARIS. *Phil. Mag.* **13**, 99 (1932).
(148) —— *Phil. Mag.* **16**, 50, 61 (1933).
(149) KEMP. *Proc. phys. Soc.* **44**, 151 (1932).
(150) OBATA, MORITA and YOSIDA. *Rep. aero. Res. Inst. Tokyo*, **79** (1932).
(151) —— —— —— *Rep. aero. Res. Inst. Tokyo*, **80**, 389 (1932).
(152) —— —— —— *Rep. aero. Res. Inst. Tokyo*, **99** (1933).
(153) HUBBARD. *J. acoust. Soc. Amer.* **4**, 138 (1932).
(154) DORSEY. *J. acoust. Soc. Amer.* **3**, 428 (1932).
(155) EISNER and KRUGER. *Hochfrequenztech. u. Elektroakust.* **42**, 64 (1933).
(156) FLORISSON. *C.R. Acad. Sci.*, Paris, **194**, 1149 (1932).

SPECTROSCOPY

§ 1. Band spectra, by R. W. B. Pearse, A.R.C.S., Ph.D.
§ 2. Infra-red photography, by H. Dingle, A.R.C.S., D.Sc.
§ 3. Grating theory, by H. Dingle, A.R.C.S., D.Sc.
§ 4. Ultra-violet spectroscopy and term-analysis of spectra, by H. Dingle, A.R.C.S., D.Sc.
§ 5 Microphotometry, by H. Dingle, A.R.C.S., D.Sc.
§ 6. Absorption and fluorescence spectra of crystals, by H. Dingle, A.R.C.S., D.Sc.
§ 7. The hyperfine structure of spectrum lines, by E. Gwynne Jones, Ph.D.
§ 8. Historical notes on wave-length standards, by C. V. Jackson.
§ 9. More recent progress in wave-length standards, by C. V. Jackson.
§ 10. Stellar research, by A. Hunter, A.R.C.S., B.Sc.
§ 11. Solar research, by A. Hunter, A.R.C.S., B.Sc.
§ 12. Instrumental technique in astrophysical spectroscopy, by A. Hunter, A.R.C.S., B.Sc.

§ 1. BAND SPECTRA

Spectra of diatomic molecules. Progress in the analysis of the complicated spectra known as "band" spectra and in the application of information obtained from them has been very great since the time when this name was given to distinguish such close, often unresolved, structures from the clearly defined lines. The first important step forward was taken when it was realized, from accumulated experimental evidence, that band spectra were emitted by molecules whereas line spectra were emitted by single atoms.

The earliest and simplest form of research in this field is that which consists in recording, by map or by measurements, the spectra obtained under various conditions from various sources and in determining the molecules responsible for the band structures observed. The results obtained in this way from systematic laboratory experiments may then be applied to determine the molecules present in new or distant sources in a manner similar to that used to detect atoms from a study of line spectra. This, of course, has considerable importance in the field of astrophysics. Examples such as the discovery of C_2, CH and CN in comet-heads, of N_2^+ and CO^+ in comet-tails and of various hydrides such as CH, OH, MgH in sunspots, are well known. This type of research is being actively pursued at the present time in many other fields, such as combustion, photochemistry, and collision processes, where studies of the emission and absorption spectra often yield valuable information as to the compounds present during various stages of the reactions. This is particularly true of such compounds as CH, NH and OH, which though stable in the physical sense, are yet too reactive to be stable in the chemical sense. Such applications have increased very greatly during the last few years, especially in connexion with photochemistry; their success depends on the amount and the completeness of the material obtained by the band-spectroscopists, and on its accessibility.

To obtain further information from a band spectrum it is necessary to analyse its structure in greater detail. We shall see how each step forward in the understanding of the structure of the spectrum increases the scope of its later application. Deslandres first made progress in this direction when he discovered that the wave-numbers of the sharp edges called "band-heads," as well as the wave-numbers of the lines into which the bands were resolved with high dispersion, could be represented by simple parabolic formulae in which certain parameters were given integral values, but little further significance could be attached to these empirical formulae before the application of the ideas of the quantum theory.

Application of quantum-theory ideas to molecular spectra met with immediate success. According to this theory, the energy of atoms or molecules can have only certain discrete values, which have been called "energy levels" or "states." Emission or absorption of radiant energy by the atom or molecule corresponds to passage from one state to another. The frequency $c\nu$ of the radiation is connected with the energies E', E'' of the states by the equation $c\nu$ E', E''

$$h.c\nu = E' - E'',$$

where c is the velocity of light *in vacuo*, and h is Planck's constant. c, h

Calculation of the energy values required the quantization of some simple model of a molecule. Quantization of the angular momentum of a rigid rotator led to a simple formula for energy E_r of rotation E_r

$$E_r = \frac{h^2}{8\pi^2 I} N^2 \qquad [N = 0, 1, 2, \ldots],$$

where I is the moment of inertia of the molecule, while quantization of a simple anharmonic motion led to the formula for vibrational energy E_v, I E_v

$$E_v = \omega v - \omega x v^2 \ldots \qquad [v = 0, 1, 2, \ldots],$$

where ω is the frequency of vibrations of very small amplitude, and x is a small fraction. The total energy of the molecule was then regarded somewhat arbitrarily as made up of three independent parts, electronic energy, vibrational energy and rotational energy, thus ω, x

$$E = E_e + E_v + E_r. \qquad E,\ E_e$$

With the guidance of these formulae the empirical analysis of molecular spectra made rapid progress. It was found that Deslandres's expressions for band-heads and band-lines could be rearranged into a form consistent with these expressions for vibrational energy and rotational energy. This was an extremely important advance, as it meant that observations of the spacing of bands and band-lines in the spectrum could be made to yield values of such molecular constants as the moment of inertia I, the internuclear distance r, and the frequency of vibration of the nuclei ω. Further r relations between these constants suggested by the laws of mechanics could be tested.

Another important result of the application of these quantum-theory formulae was the unification afforded by the linking of infra-red observations with the visible and ultra-violet spectra. Far infra-red spectra were found to be due to changes of rotational energy only, near infra-red spectra to changes of vibrational energy and rotational energy simultaneously, and visible and ultra-violet spectra to simul-

taneous changes of electronic, vibrational and rotational energy. Verification lay in the finding of identical values of molecular constants from all three portions of a spectrum—in finding, for example, the same value for the moment of inertia from rotation bands, from the lowest state for vibration-rotation bands and from the lowest state for electronic bands. Such observations have been made in the case of the spectra of the hydrogen halides.

Much has been learnt from the study of the spectra of molecules formed from atoms possessing isotopes. For all isotopes of any one element the electronic configurations are the same, so that in the molecules we may expect the interatomic forces to be closely the same. But the masses of the nuclei are different, so that the moments of inertia and frequencies of vibration are also different. The formulae show that this causes differences in the energy of the rotational and vibrational states leading to shifts of the lines and bands of the spectrum. Detection and study of these shifts has afforded independent evidence in support of the findings of the mass spectrograph, and has led to the discovery of such rare isotopes as O^{18}, O^{17}, N^{15} and C^{13}. "In some cases, as in $B^{11}O$ and $B^{10}O$, the mass-ratio has been obtained with a high degree of precision. On the theoretical side, study of the isotope displacements has been important in deciding between energy values given by the various forms of the quantum theory: the necessity for half-integral quantum numbers is an example. Recently, the study has received a great stimulus from the discovery of an isotope of hydrogen. The large mass-ratio, 2, leads to shifts in the spectrum which are relatively enormous, so that the behaviour of higher terms in the energy formulae may be observed with some accuracy.

The application to chemistry of the detailed knowledge of molecules obtained from band spectra, which promises to have the most far-reaching results, was begun when Birge and Sponer showed the relation between the equation for the vibrational energy-levels of a molecule and the heat of dissociation. As the vibrational energy of a molecule in a particular electronic state is increased by addition of successive quanta, the spacing between the levels decreases until they merge into a continuum. Extrapolation of the vibrational energy formula to this point yields a value for the energy required to dissociate a molecule in this particular electronic state. In this connexion the curves, first proposed by Franck, showing the relation between the potential energy of the molecule in its various electronic states and the internuclear distance have been very valuable in correlating spectroscopic facts and in picturing the mechanism of dissociation.

So far only the positions and spacing of bands and lines have been considered, but much may be learnt in addition by study of the intensities. Besides depending on the statistical weights and transition-probabilities of the levels concerned, the intensities depend on the number of molecules emitting or absorbing and on their distribution among the levels—among the lower levels for absorption and among the upper levels for emission. Measurement of the intensities thus leads to information about the distribution for various conditions of excitation; in the case where there is thermal equilibrium, application of the Boltzmann factor makes possible an estimate of the temperature of the source.

The results so far described followed from the application of the quantum theory to simple models of the molecule. Despite the progress made, a great deal remained to be explained in regard to the details of the structure of bands. Bands showed multiple electronic levels—singlets, doublets, triplets; sometimes the structure of a band fell into three main branches—*P*, *Q* and *R*—corresponding to changes of $+1$, 0 and -1 respectively, in the rotational quantum number, and sometimes only into two, *P* and *R*, corresponding to changes $+1$ and -1; occasionally bands showed other fainter satellite branches; there were the problems of the missing lines and the combination defect between the *Q* branch and the *P* and *R* branches. Evidently greater knowledge of the nature of the electronic states of the molecule was required.

A first step in the understanding of the electronic states was the derivation and application of the Kramers-Pauli formula, which took account of a resultant electronic angular momentum inclined at a fixed angle to the internuclear axis. This was of considerable help in deriving empirical formulae to express the energy states found in many bands, but left much uncertainty in the assignment of values to certain of the quantum numbers involved. The foundation of the present well-established theory was laid by Hund when he applied to molecules the ideas of the electron-spin vector model which had proved so successful in ordering line spectra. This he did by considering the reactions between the various angular momenta and defining certain limiting cases according to the relative strengths of these interactions. Of these cases Hund's case *a* and case *b* have found widest application.

With this model as basis a standard notation has been evolved, chiefly through the exertions of Hund, Mulliken, Wigner and Witmer. This has been devised to exhibit as clearly as possible the analogies which exist between atomic electronic levels and molecular electronic levels, and to simplify the correlation of one with the other. With the aid of the methods of the newer wave mechanics, this correlation is rapidly paving the way to a new and more detailed chemistry—a chemistry of the interaction of individual atoms and molecules.

Recent results of wave mechanics have also shed light on other phenomena exhibited by molecular spectra. Among these may be mentioned predissociation. In some cases this is shown by the rotational structure of certain bands becoming diffuse, in others by sudden breaking off of the branches. Kronig has explained this as due to the finite probability, under certain conditions, of a molecule in an excited state undergoing a radiationless transition to another state in which its energy is sufficient to cause dissociation. Another interesting phenomenon is the alternating intensities of lines of the branches of bands emitted by homonuclear molecules. The ratio of the intensities depends on the value of the nuclear spin, so that photometric measurement of this ratio in band spectra has been able to supplement in a valuable way determinations of nuclear spin from observation of hyperfine structure of atomic spectra. This is particularly true in the case of light atoms.

Spectra of polyatomic molecules. The analysis of the spectra of polyatomic molecules has not reached nearly such a degree of completeness as that of diatomic molecules; it is in fact just at its beginning. The reason for this is the large number of

difficulties which have to be overcome both on the experimental and on the theoretical sides. In the first place, the structure is so closely spaced that very high dispersion is required, yet at the same time it is almost impossible to use intense sources of emission without dissociating the molecules. Thus only in absorption can the spectra be observed conveniently. Unfortunately the majority of the absorption spectra are in the extreme ultra-violet, where high dispersion presents greater difficulty. In the second place, when the spectrum is obtained its complexity is usually such as to defy analysis. Whereas for each electronic state of the diatomic molecule there is one principal frequency and one moment of inertia, in a polyatomic molecule of n atoms there are three principal moments of inertia and $3n-6$ proper frequencies. In the third place, except for simplified cases, the mathematical treatment is not easy and is far from complete. However, in spite of all these difficulties very considerable progress has been made. This has in the main taken the form of systematic study of the vibration frequencies observed in series of chemically related compounds. In this way it has been possible to ascribe certain frequencies to particular bonds such as O—H, N—H and C—H, and also to calculate and explain to some extent the proper frequencies of certain molecules. This work has been much aided by observations of the Raman spectra, which give independent values for certain of the fundamental vibration frequencies. Analysis of rotational structure presents very great difficulties on account of its complexity, but in a few cases a good start has been made. Examples are the analyses of the absorption spectra of H_2O and C_2H_2 by Mecke and Hedfeld, of the absorption spectrum of HCHO by Dieke and Kistiakowsky, and of the emission spectrum of CO_2 by Schmid.

§ 2. INFRA-RED PHOTOGRAPHY

One of the most important developments of recent times is the commercial manufacture of photographic plates sensitive for a considerable distance into the infra-red. This has extended the possibilities of research in all branches of spectroscopy. Although the exposures necessary are as yet considerably longer than for the ordinary regions, and the range of infra-red sensitivity of a single plate is scarcely more than 1000 A., the advance over previous possibilities is nevertheless enormous, and the photographs when obtained are hardly inferior in quality to those given by ordinary plates. Wave-lengths up to 12,000 A. or beyond can now be photographed without much difficulty if a bright source is used. The routine work of recording and measuring the wave-lengths of the spectra of various elements in the newly opened regions has been undertaken chiefly at the Bureau of Standards in Washington, although contributions have been made by workers elsewhere. Accurate standards of wave-length remain to be determined, but it should be noted that for a great deal of spectroscopic work—notably the analysis of spectra—in which wave-numbers are used, the error in wave-number corresponding to a given error in measured wave-length is so small in comparison with its value in the more familiar regions, that temporary standards of a lower order of accuracy than that necessary for the highest precision are sufficient.

§ 3. GRATING THEORY

At the other extreme of the spectrum—the region of very short ultra-violet waves—work of great interest and importance has been carried out at Upsala by the use of plane and concave gratings with light at grazing incidence. Although the general principles underlying the properties of spectrographs at grazing incidence have long been known, it is only recently that Siegbahn has shown how they can be applied with great advantage in the extreme ultra-violet. The chief advantages arise from the increase of reflecting-power and resolving-power with angle of incidence. The improvement in reflecting-power is specially marked with short wave-lengths, and even a slight increase of angle of incidence, when this angle is large, results in a considerable extension of the spectrum towards the shorter wave-lengths.

The question of resolving-power calls for more detailed consideration than it has hitherto received and this has been carried out for the concave grating by Mack, Stehn and Edlén[1], who have extended the existing theory to the fourth order approximation. In the course of the work it was found desirable to make a slight modification of Rayleigh's well-known criterion for resolution of two lines of equal intensity (namely, that the lines are just resolved when the central maximum of one coincides with the first minimum of the other). A more satisfactory criterion appears to be that two such lines can just be resolved when, in the combined intensity curve, the intensity at the central minimum is $8/\pi^2$ (i.e. 0·8106) as great as that at either of the chief maxima.

With this criterion the resolving-power can be computed in terms of the *effective width* of the grating; i.e. the width of a plane grating having the same resolving-power as the concave grating in question. The effective width is not identical with the actual width of the grating, because the customary expression for resolving power (mN, where m is the order of spectrum and N is the number of lines ruled on the grating) is in the new approximation multiplied by a factor which varies with grating width, and when the latter exceeds a certain optimum value this factor reduces the resolving-power in spite of the increased number of lines. Mack, Stehn and Edlén have calculated the optimum widths for a 1-metre grating with various values of the wave-length, grating space, order of spectrum, and angle of incidence, and have shown that for large angles of incidence and short wave-lengths the gratings ordinarily used in the past have been considerably wider than the optimum, so that resolving-power has actually been lost through the use of too large a grating-aperture. The relation of spectrum-intensity to grating-width has been investigated also, and an optimum width for this property, having a value practically identical with that for resolving-power, is found to exist. m, N

Not only, however, is an increase of grating-width beyond this value attended by the loss of desirable features in a spectrum, but at large angles of incidence it introduces also an undesirable feature in the form of scattered light which appears as ghosts. These ghosts are not identical with the familiar Rowland ghosts obtained under the usual conditions—they have, for example, no such simple regularity of spacing—but their reality is convincingly shown by photographs taken under

identical conditions with various widths of the exposed portion of the grating. It appears, therefore, to be a matter of considerable importance to design concave-grating spectrographs in such a way that they are readily adaptable to the region of spectrum and angle of incidence to be used. The optimum width for an angle of incidence of 88°, for example, varies from 0·53 cm. at λ 10 A. to 2·86 cm. at λ 5000 A. for the 1st order of a 1-metre grating having 6000 lines per cm., and these values are scarcely affected by a four-fold increase in the order of spectrum or in the number of lines per centimetre of the grating. At normal incidence, on the other hand, such a change is sufficient to halve the optimum width. The conclusion that, for vacuum work at grazing incidence, the grating aperture should not exceed 1 cm. or less is very surprising, and, from the point of view of grating manufacture, very gratifying also.

§ 4. ULTRA-VIOLET SPECTROSCOPY AND TERM-ANALYSIS OF SPECTRA

From among the excellent researches which have been carried out by the new method we may select for special notice an investigation by Edlén[2] into the several spectra emitted by the six lightest elements next to hydrogen and helium—lithium, beryllium, boron, carbon, nitrogen and oxygen. As is well known, the number of possible spectra obtainable from an element is equal to its atomic number. These
Z spectra (for a hypothetical element, *Z*) are denoted by *Z* I, *Z* II, ..., beginning with the spectrum of the neutral atom and proceeding through the successive stages of ionization down to the hydrogen-like ion with only a single electron. In the early days of modern atomic theory, complete sets of spectra were known for two elements, hydrogen and helium, with atomic numbers 1 and 2 respectively; but for many years it was impossible to add to this number, partly on account of the difficulty of supplying sufficient energy to bring about the successive degrees of ionization, and partly because the technique of ultra-violet spectroscopy (at each stage of ionization the spectrum of an element recedes farther into the ultra-violet) had not been adequately mastered. By the grazing-incidence method, Edlén has now acquired sufficient data to enable the main features of all the spectra possible for the first eight elements of the periodic table to be elucidated. As a source of light he has employed a high-tension vacuum spark, passing between poles containing the substance under investigation. Adjustable self-inductance in the circuit enabled the degree of ionization to be controlled, so that the successive spectra of each substance could be distinguished from one another.

The determination of standard wave-lengths in a region where the interferometer cannot, at least at present, be employed for this purpose was carried out by a comparison of the 1st-order spectrum of the Pfund iron arc, which has been measured interferometrically, with certain high-order lines of oxygen, carbon and aluminium. As an example of the high orders reached, Edlén states that the 17th order of λ 160 (Al IV) and the 20th order of λ 100 (Be III) were obtained with good intensity and definition. A number of standards were measured in this way to the nearest thousandth of an angstrom, and other lines were then measured by interpolation, a correction curve of the ordinary type being used.

For each of the spectra studied a table of wave-lengths is given, and the excellence of the spectrograms, many of which are reproduced, is shown, for example, by the fact that components of the C IV group at λ 238, separated by 0·050 A., are resolved and measured separately. The importance of obtaining the highest possible resolution in this region will be realized when it is stated that this small difference corresponds to a separation in wave-number of 88 $cm.^{-1}$ Edlén considers the various spectra obtained in terms of spectrum theory, and gives the term-analyses as far as possible, accompanied in many cases by a Grotrian diagram. In a valuable account of many details of the spectra he discusses such questions as intensity relations, the suitability of certain lines as standards for wave-length determinations, and the relation of the results to the requirements of quantum mechanics and to the semi-empirical regular and irregular doublet laws. In conclusion, a general table shows the values obtained for the deepest terms of all the spectra theoretically possible for the first eight elements, together with the ionization potentials, which range from 5·364 V. for Li I to 867·087 V. for O VIII.

The rapid progress in the analysis of spectra which has taken place during the last few years has shown the need for a collection of data on this subject. Such a collection is desirable to prevent excessive repetition, to facilitate the use of known term-values in several atomic problems, and to allow of a more general comparison of the characteristics of spectra of the various elements and ions than has so far been possible. This need has now been supplied by the publication of *Atomic Energy States* by R. F. Bacher and S. Goudsmit (McGraw Hill Co., 1932). This valuable work gives, for each element and ion for which data were available at the time of publication, the ionization potential, outer electron structures, and corresponding term-values. It has been well prepared and indexed, and although such a work must necessarily require continual supplementing and revision as knowledge accumulates, the greater part will be of lasting value.

§ 5. MICROPHOTOMETRY

A fruitful subject of study in the immediate future will be the contours of spectrum lines in relation to the physical conditions and the number of absorbing or emitting atoms in the spectrum source. This has been made possible by the development of microphotometry, and there are several theoretical problems awaiting solution which can best be approached in this way. It would scarcely be profitable at this stage to record what has been done, as it is mostly of a preliminary character; and although it is of great value in the development of the subject, the investigators cannot yet be said to have established a standard procedure either in method of investigation or in interpretation of results. Sufficient experience has been obtained, however, to show that both the practical and theoretical problems involved are by no means so simple as they might at first appear to be.

§ 6. ABSORPTION AND FLUORESCENCE SPECTRA OF CRYSTALS

Another subject whose inclusion in such a report as this is justified more by promise than by present achievement, is that of the absorption and fluorescence spectra of crystals. In many instances it has been found that such spectra are due not to the main constituent of the crystals but to impurities, often present in very small quantity. Certain metals of the rare earths, particularly samarium and europium, are specially active in this way, and the phenomenon is interesting, among other reasons, because it affords one of the most marked distinctions known between elements which chemically are so strikingly similar. A very curious feature of the phenomenon is that in some examples of fluorescence the line spectrum of the metal concerned is shown, although the substance is presumably in the solid state. A study of the lines which appear in relation to the electronic structure of the atoms should afford valuable information. Among others, Przibram, Deutschbein, and Spedding with several collaborators have made valuable investigations of the spectra in connexion with the general problem of crystal structure, and have thrown much light on its chemical as well as its physical aspects. As an example of the variety of ways in which such phenomena, which at one time were considered in purely chemical aspects, can now be studied, may be mentioned also the work of Du Bois and Elias on the effect of a magnetic field on the line fluorescence of the ruby. One of the chief desiderata would appear to be the employment of greater dispersion in many of the investigations, but this is made difficult by the extreme faintness of much of the fluorescence. It is sufficiently evident, however, that an important field of inquiry here lies open to experimental investigation.

§ 7. THE HYPERFINE STRUCTURE OF SPECTRUM LINES

The production of large diffraction gratings and the invention of the various forms of interferometers first made it possible to study spectrum lines in detail under high resolving-power. It was at once apparent that such lines were rarely simple and symmetrical as was to be expected from the kinetic theory, and early in this century a considerable literature on line structures had accumulated. The data so presented were not, however, very satisfactory, as there was a great deal of disagreement between the results obtained by the various workers. It is now generally recognized that this variability was conditioned by the unsuitable light-sources and imperfect instruments used at that time. Researches by Schüler have since demonstrated that in suitable discharge tubes the lines of elements of all classes may be excited free from self-reversal or accidental modifications of structure, and the increased use of apparatus of better design has largely eliminated the uncertainties due to instrumental ghosts. As soon as it had become possible to obtain reliable data, progress in the analysis and interpretation of the line structures was very rapid.

Experimental details. The mode of excitation of the spectrum naturally depends upon the type of element and on its atomic weight and vapour pressure at ordinary temperatures; if the lines involve low terms in the spectrum, special precautions must be taken to reduce the amount of vapour between the luminous region and the

window of the discharge tube, or else serious errors due to self-reversal may be introduced. Some of the heaviest elements may be excited in an arc if reasonable care is taken, but this source has now fallen out of favour on account of its many disadvantages. Most observers use a form of hollow-cathode discharge tube, or "Schüler lamp," which is particularly valuable because, by a proper choice of excitation conditions, arc or spark lines can be produced at will. The ordinary Geissler tube and the high-frequency discharge are often employed to excite the lines of gaseous or easily vaporizable elements. The importance of the light-source cannot be too strongly emphasized, for apparatus of high resolving-power is useless if the line structures are distorted in the discharge tube.

The Fabry-Perot interferometer with silvered mirrors is very widely used for examining line structures. It has many advantages over the other types of interferometric apparatus (Lummer plates and echelon gratings) in the visible and infra-red portions of the spectrum, but its resolving power is seriously diminished in the ultra-violet unless special mirrors can be obtained. The new reflection echelons promise to yield interesting data in the violet and ultra-violet regions, but they resemble the Lummer plates in that their dispersion is not variable as in the Fabry-Perot interferometer, so the instrument must be crossed with another interferometer in order to remove ambiguity due to the overlapping of orders. Large gratings are, of course, still employed in special cases where the structures are wide.

The apparently capricious appearance of hyperfine structures and the great complexity of certain line patterns made it obvious that, in general, the phenomenon was not due simply to a separation of the lines arising from the various isotopes. On the other hand, the first hyperfine structures to be explained were those occurring in the neon spectrum and these are all due to isotopic shifts, so that some, at least, of the patterns in other elements may be explicable in this manner. Many of the neon lines consist of narrow doublets, the component on the short-wave side having about a tenth of the intensity of the main component in each case. Since neon has two isotopes (20 and 22, the third, 21, being too faint to concern us here) with just this abundance-ratio, the proposed isotope shift explains the structures completely. More recently, doublet structures found in the arc spectra of potassium (isotopes 39 and 41), chlorine (35 and 37), lithium (6 and 7) and hydrogen (1 and 2) have been traced to a similar isotope effect. It should be pointed out that if the hyperfine structures of the lines of any element are due solely to isotope shifts, all the patterns must be alike in regard to number, relative intensity and arrangement of components, though the spacing may vary slightly in lines that involve different terms. The hyperfine structures exhibited by most elements, bismuth and mercury being good examples, vary so much from line to line that it is clear that a fundamentally different explanation must be given in such cases.

Nuclear moments. In 1924, Pauli suggested that if the nucleus of an atom were endowed with mechanical and magnetic moments, each line in its spectrum should have a hyperfine structure resembling that of an ordinary multiplet; hence the patterns should be analysable by the methods that are used in multiplet-analysis. Although various authors had tried to derive level-schemes from hyperfine-structure

data, no great success was obtained until the classic work by Goudsmit and Back on the bismuth arc spectrum was published. This investigation showed that the structures of a number of Bi I lines could be attributed to hyperfine structures in certain of the energy-levels and hence proved that the bismuth nucleus does indeed possess
I a spin moment I which is coupled with the outer electrons of the atom through a nuclear magnetic moment μ. In view of the importance of this work, we will discuss

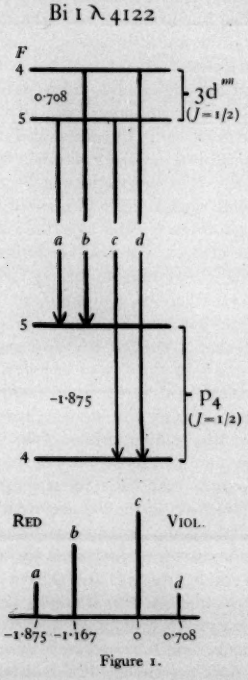


Figure 1.

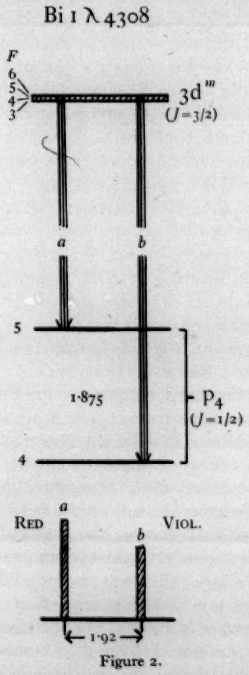


Figure 2.

the analysis of a few of the lines in order to give an insight into the methods used and the results obtained.

λ 4122, which consists of two pairs of lines, each with the same separation
$\Delta\nu$ $\Delta\nu$, equal to 0·708 cm.$^{-1}$, is at once recognized (by comparison with patterns found in multiplet spectra) as arising from a combination between two levels each having a narrow doublet structure. Figure 1 shows that only one interpretation is possible: the two levels have hyperfine separations of 1·875 and 0·708 cm.$^{-1}$ respectively, but it is not possible at this stage to know which term has the larger separation. We now

turn to other lines involving (say) the p_4 level, and notice that λλ 4493, 4308 and 3888 all have doublet structures with separations varying from 1·81 to 1·92 cm^{-1}. In some cases the components are rather diffuse, indicating that the patterns are not completely resolved, so it is evident that the separation 1·875 cm^{-1} must be assigned to p_4 whilst the three other levels (3d′, 3d‴ and 3s respectively) have structures that are too narrow to be resolved by the grating employed. Contrast the structures of λ 4308 and λ 3596 in figures 2 and 3. Returning to λ 4122, we see that the separation 0·708 is to be assigned to 3d⁗. A further check on these conclusions is obtained from the analysis of other lines, such as λ 3596. In figure 3 we see that the differences *a*–*e* and *b*–*f* are 1·875, which gives the key to the whole pattern. The structure of the upper term, 4d″, is of importance since it leads to the first allocation of quantum-numbers F from a consideration of the interval rule, in which we seek to assign F values to each of the sub-levels so that the factor $(\nu_F - \nu_{F-1})/F$ shall be the same for all the intervals. This factor is 0·095 cm^{-1} for 4d″, and the F values are 6, 5, 4 and 3. The selection rule ($\Delta F = \pm 1$ or 0) requires hyperfine quantum-numbers 5 and 4 for the terms p_4 and 3d⁗.

F

The hyperfine structures of these Bi I terms are interesting as they show how the term multiplicities may be limited by the numerical value of J, if this is less than I. It is a generally valid rule that the maximum number of sub-levels into which a term may be split by a nuclear moment is $(2N+1)$ where $N = J$ or I, whichever is the smaller. Thus, by inspection it is clear that $I > \frac{9}{2}$ for bismuth. We see from the F values of the term 4d″ that I is probably $\frac{9}{2}$—a value which is confirmed by a study of the Zeeman effects shown by the individual hyperfine structure components; but since this phenomenon can only be studied in a few exceptional cases such as bismuth, thallium and rhenium, where the structures are wide, it will not be discussed here.

N

The method of analysis so described has been completely successful in the interpretation of the hyperfine structures found in the lines of the non-isotopic elements with odd atomic weights and of those elements which consist of odd isotopes alone (e.g. Al, As, Au, Bi, Br, Co, Cs, Cu, F, I, In, K, La, Mn, Na, Pr, Re, Ta, Tl and V). These investigations have shown that the arrangement and number of components in a hyperfine structure are determined by the value of the nuclear spin I, whilst the total width of the pattern is governed by the magnitude of the nuclear magnetic moment μ. In addition, interval and intensity rules resembling those previously known for multiplet spectra have been formulated and experimentally tested.

Elements with odd and even isotopes. On the other hand, structures found in the lines of elements containing both odd and even isotopes have led to very different results when analysed in the usual manner, so that either gross violations of the interval and intensity rules must be accepted or a completely new explanation must be sought. The way out of the difficulty was discovered by Schüler and Brück who, working on the spectrum of cadmium (isotopes 110, 112, 113, 114, 115 and 116), proved that only the lines of the odd isotopes show the usual hyperfine structures, whilst the lines of the even isotopes are not split by the action of nuclear spins and

are thus superposed upon the ordinary hyperfine pattern of the odd isotopes. Hence, in such cases, the otherwise explicable hyperfine multiplet is complicated or masked by the presence of an unresolved line—much more intense than any of the other components, since the even isotopes are normally more abundant than those with odd mass-numbers—which falls at the optical centre of gravity of the pattern. After

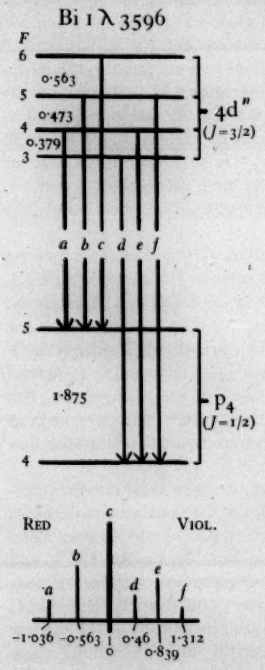


Figure 3.

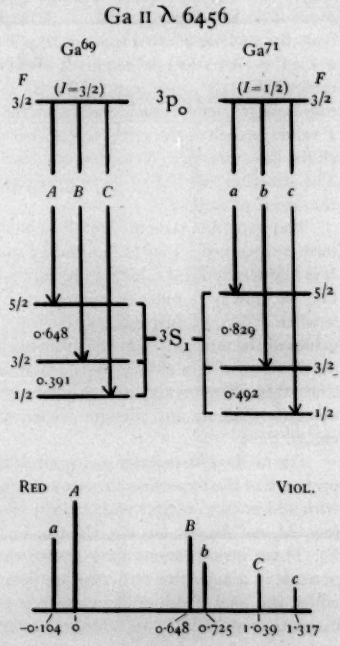


Figure 4.

removal of this foreign line the structure may be analysed in the ordinary way. The hyperfine structures in barium, krypton, lead, tin, strontium and zinc have since been investigated with similar results.

When the element concerned has two odd isotopes, the two hyperfine patterns are usually identical and coincident, so the analysis proceeds as if there were only one present. This means that the odd isotopes of the given atom have the same nuclear spins and magnetic moments. In the lines of gallium, mercury, rubidium, antimony and xenon, however, we find, apart from the lines due to any even isotopes that may

be present, two independent hyperfine patterns of different total widths, indicating that here the nuclear magnetic moments have different magnitudes. Whilst in gallium and antimony the two patterns are similar, figure 4, in mercury, rubidium and xenon the patterns differ also in regard to number and arrangement of components, figure 5, proving that in these three elements the nuclear spins of the two odd isotopes are no longer the same. With few exceptions (which will be discussed

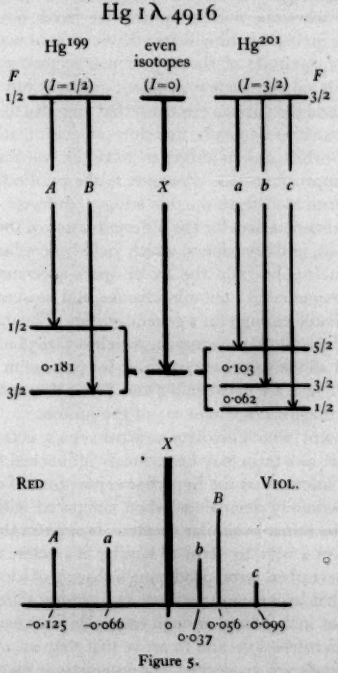


Figure 5.

later) the centres of gravity of both patterns fall together and coincide with the lines arising from the even isotopes, if such be present.

In a few elements, notably mercury, thallium and lead, some of the line structures are still not completely explained by the above scheme. These structures, which form the minority in all the spectra concerned, are complicated by the appearance of isotopic shifts. The hyperfine multiplets of the odd isotopes may still be traced to the action of nuclear spins, but the centres of gravity no longer coincide. In addition, although the lines arising from the even isotopes individually show no hyperfine

structures they do not fall together, but appear as a number of equally spaced components in the midst of the line pattern. The spacing of the lines due to the individual even isotopes is equal to the separation of the centres of gravity of the odd isotopes, and their relative intensities are in agreement with the known abundance ratios.

The principal advances in the year under review have been in the calculation of nuclear magnetic moments μ. The necessary experimental data for such calculations are usually provided when the nuclear spins I are fixed, but the previous results cannot be regarded as giving anything better than orders of magnitude except in the simplest cases. The magnitude of the hyperfine-structure separations in a given term depends not only on the nuclear magnetic moment but also on the coupling between the nucleus and the various electrons that give rise to the energy term. An evaluation of these quantities demands, therefore, an exact mathematical description of the outer electron-orbits and an ability to solve the resulting equations without the use of doubtful approximations. Previous to the publication of two papers by Goudsmit and by Fermi and Segrè on this subject, different values of μ were obtained from the experimental data for the different terms of the same spectrum, but formulae are now given, in these papers, which yield concordant values whether the data for the given nucleus be from the arc or spark spectrum, or from *s*, *p*, or *d* terms. This must be regarded as a notable advance, and most of the values in table 1 would seem to be reliable enough for a general survey. The magnetic moments are expressed in terms of the nuclear magneton, which is 1/1838 of the Bohr magneton. It will be noticed that all the values except those for potassium and silver, which are much smaller than the rest, are between 0·5 and 5·5 nuclear magnetons and are thus comparable with the magnetic moment 2·5 of the proton.

Another development, which involves new principles, is the demonstration that the hyperfine structure of a term may be seriously influenced by other terms in the same spectrum. It is known that the hyperfine separations of a number of levels in certain spectra are definitely anomalous when compared with those of the other terms and of analogous terms in similar spectra. It appears that the cause of these perturbations is always a near-by term of similar character, which engages in an interaction with the perturbed term, producing a sharing of identities and properties between the two so that hyperfine structures and isotope shifts are greatly altered. In spite of the rarity of such effects, enough examples have been found to show the main features of the perturbations and to prove that they are objective phenomena. As far as hyperfine levels are concerned, the interactions take place only between levels of the same F value and, if all the sub-levels are affected (i.e. if both interacting terms have the same J value), the net result is to alter the interval factor. When the J values differ, however, only a few of the hyperfine levels are affected and serious deviations from the Landé interval rule occur. The latter effect can be observed only when the terms are very close together, and only one case has yet been reported.

Table 1. Nuclear mechanical and magnetic moments

Atomic number Z	Symbol	Mass numbers	Nuclear spin I*	Nuclear magnetic moment μ
1	H	1	1/2 (B)	2·5
		2	1 (B)	0·8
2	He	4	0 (B)	—
3	Li	6	0	0
		7	3/2	3·29
6	C	12	0 (B)	—
7	N	14	1 (B)	—
8	O	16	0 (B)	—
9	F	19	1/2	—
11	Na	23	3/2	2·0
13	Al	27	1/2	2·1
15	P	31	1/2 (B)	—
16	S	32	0 (B)	—
17	Cl	35	5/2 (B)	—
19	K	39	(?)5/2 (B)	—
23	V	51	5/2	—
25	Mn	55	7/2	—
27	Co	59	9/2	—
29	Cu	63, 65	3/2	2·5
30	Zn	66, 68, 70	0	0
		67	3/2	—
31	Ga	69	3/2	2·01
		71	3/2	2·55
33	As	75	3/2	0·9
35	Br	79, 81	3/2	—
36	Kr	78 to 86	0	—
		83	7/2 or 9/2	—
37	Rb	85	5/2	1·3
		87	3/2	2·7
38	Sr	86, 88	0	0
		87	3/2	1·0
47	Ag	109, 111	?	<0·25
48	Cd	110 to 116	0	0
		113, 115	1/2	−0·6
49	In	115	9/2	5·4
50	Sn	116 to 124	0	0
		117, 119	1/2	0·9
51	Sb	121	5/2	2·7
		123	5/2	2·1
53	I	127	5/2	—
54	Xe	124 to 136	0	—
		129	1/2	—
		131	3/2	—
55	Cs	133	7/2	2·6
56	Ba	136, 138	0	0
		137	3/2 (?)	0·9
57	La	139	7/2	—
59	Pr	141	5/2	—
73	Ta	181	7/2	—
75	Re	187, 189	5/2	—
79	Au	197	3/2	1·8
80	Hg	198 to 204	0	0
		199	1/2	0·55
		201	3/2	−0·62
81	Tl	203, 205	1/2	1·8
82	Pb	204, 206, 208	0	0
		207	1/2	0·60
83	Bi	209	9/2	0·89

* (B) denotes that the mechanical moment in question is determined solely from measurements of band spectra.

§ 8. HISTORICAL NOTES ON WAVE-LENGTH STANDARDS

The unit of length used for the measurement of wave-lengths is called the *angstrom*, after the Norwegian physicist A. J. Ångström, who, with the aid of three gratings ruled by Novart, first made a map of the solar spectrum with a wave-length scale instead of an arbitrary scale. The unit was defined as 10^{-10} m., and was sometimes in the older spectroscopic literature known as the "tenth-metre."

Ångström's map (1868) was subsequently displaced by Rowland's table of solar wave-lengths, which for many years was used as a standard of comparison for the determination of wave-lengths. This system of solar standards was made by the use of concave gratings which Rowland invented specially for the purpose, and was thought by Rowland to have a relative accuracy of 0·01 A. Rowland made no attempt to measure the wave-length of any of the solar lines absolutely, but assumed that, for D_1 of Na, $\lambda = 5896{\cdot}156$ A., at 760 mm. pressure and 20° C., and determined all the other wave-lengths relatively to this standard. His method was dependent on the accuracy of the principle of overlapping orders.

At about the same time (1888) Kayser and Runge gave a map and wave-length table of the iron arc spectrum on the same scale as Rowland's solar table and of similar accuracy.

Soon after the publication of Rowland's work, A. A. Michelson applied the interferometer to the study of spectrum lines, and by 1893 had determined the wave-lengths of the red, green and blue lines of cadmium by direct comparison with the metre at Sèvres. These wave-lengths were nearly a hundred times more accurate than absolute values found by the grating method, and showed that the wave-length adopted by Rowland for D_1 was too high by about one part in 30,000.

However, this affected only the absolute value of Rowland's scale, and no doubt was thrown upon the relative accuracy of the wave-lengths until a few years later.

In 1896 Jewell showed beyond any reasonable doubt that solar lines and arc lines of the same element had not exactly the same wave-lengths and that the differences were as much as 0·01 A. and occasionally even more; the relative displacement of the solar lines was usually to the red. However, Jewell was never able to convince Rowland of the reality of this effect, and Rowland insisted on making no allowance for it in his *Preliminary Table of Solar Wave-lengths*.

In 1900 Kayser found that there were *relative* errors amounting to as much as 0·02 A. in Rowland's *New Table of Standard Wave-lengths* (1893) and that the distribution of standards was inadequate for accurate interpolation. Much further light was thrown on the matter two years later by Fabry and Perot, who were able to measure directly with their new interferometer the wave-lengths of 33 solar lines with an accuracy of two or three milliangstroms. They showed (i) that Rowland's absolute values were in error by the amount suggested by Michelson; (ii) that the relative values of the preliminary table were accurate, *over small ranges*, to one part in a million; and (iii) that over large ranges the relative values were in error by amounts up to 0·02 A.

This was the state of affairs when the first meeting of the newly formed Inter-

national Union for Co-operation in Solar Research took place at St Louis in 1904. At this meeting a memoir on standards of wave-length, prepared by MM. Fabry and Perot, was presented by M. Poincaré; in it these observers put forward their discoveries which have been mentioned above. Kayser suggested that it was necessary to prepare an entirely new system of standards of wave-length which should be determined interferometrically by comparison with the red line of cadmium as primary standard. Jewell on the other hand suggested that it would be more satisfactory to use a corrected Rowland scale. No formal resolution on the matter was passed at this meeting.

The next meeting of the Union took place two years later at Oxford and the following important resolutions were adopted: (i) the wave-length of a suitable spectroscopic line shall be taken as primary standard of wave-length; (ii) secondary standards are required at distances not greater than 50 A. These secondary standards should be referred to the primary standard by the interferometer method. The source of light should be obtained by means of an electric (iron) arc of from 6 to 10 A.; (iii) a committee shall be appointed to select the standards and to organize the determination of their wave-lengths in terms of the primary standard in at least two independent laboratories; (iv) the same committee shall be charged with the selection of tertiary standards which shall be at distances of from 5 to 10 A. The wave-lengths of these tertiary standards are to be determined by interpolation with the help of gratings.

The third meeting of the International Union for Co-operation in Solar Research took place at Meudon in 1907, and at this meeting Benoît, Fabry and Perot gave a full account of their determination of the wave-length of the red line of cadmium in terms of the metre. Their value $\lambda = 6438{\cdot}4696$ I.A. for the cadmium red line was adopted as the primary standard of wave-length.

At the same meeting Fabry and Buisson gave a description of the method and apparatus they used for the interferometric determination of the wave-lengths of about 120 lines in the spectrum of the iron arc, covering the range 2373 A. to 6494 A. They gave a list containing these wave-lengths and stated that they thought their values were accurate to about one part in a million.

In the three years following, Pfund and Eversheim followed the method of Fabry and Buisson for measuring the wave-lengths of iron standards, and in the region 4282 A. to 6494 A. (common to the three sets of observations) fifty lines were adopted as secondary standards, at the meeting of the International Union for Co-operation in Solar Research at Mount Wilson in 1910. It was also decided at this meeting that the symbol I.A. should be used to designate an angstrom unit on the new scale.

During the next few years progress in the determination of standards of wave-length was slow, and when the Union met again at Bonn in 1913 it was possible to add only 33 lines to the list of iron secondary standards. However a considerable amount of work was done on the tertiary interpolated standards, especially at Mount Wilson, and some interesting discoveries were made concerning the pole effect and the pressure shift in the iron arc.

After the war there were no more meetings of the International Union for Co-operation in Solar Research, but fortunately the International Astronomical Union, which was founded just after the war and had its first meeting in Rome in 1922, appointed a Commission on Standards of Wave-length, which has continued the work formerly done by the Solar Union.

At the Rome meeting of the International Astronomical Union several important recommendations were passed. The most important of these was the adoption of a fairly complete list of iron-arc wave-lengths covering the range 3370 A. to 6750 A. This consisted of: (i) the ninety secondary standards adopted previously, which were all interferometric values and which were confirmed by the interferometric determinations of St John and Babcock made at Mount Wilson in 1921 with a source comprising 1¼ millimetres at centre of 12-mm. 5-amp. iron arc; and (ii) about 300 tertiary standards based on the interferometric determinations of St John and Babcock and of Burns, Meggers and Merrill, and the grating interpolated wave-lengths of six independent observers. The agreement between the interpolated wave-lengths and the secondary standards in this list was excellent. The difference is 0·001 A. or *nil* with only one or two exceptions. This list of iron standards is now known as *the 1922 scale*.

The other important recommendation was the adoption of twenty neon lines as secondary standards. These lines had been measured by three observers at the Bureau of Standards and one observer in Germany. They have been retained as secondary standards without any alteration up to the present time. Between 1922 and 1932 their wave-lengths have been confirmed by four observers (Monk, Wallerath, Babcock and Jackson).

During the six years following the first meeting of the International Astronomical Union, much important work was done on the wave-lengths of the iron arc in air in the region covered by the 1922 scale. The chief investigations are those of Monk (1922), Meggers, Kiess and Burns (1924), and Babcock (1927). They were all carried out with the interferometer and led to the following important conclusions: (i) the iron arc must not be less than 12 mm. long and should be longer if possible. The anode should be placed below the cathode, and should consist of a bead of iron oxide supported on a rod of iron or copper not less than half an inch in diameter. The cathode should be a rod of steel 0·25 in. in diameter, cooled by a massive cylinder of brass or copper fitted close to the end. The current should not exceed 5 A., and only a horizontal central zone at right angles to the axis of the arc, not exceeding 1·5 mm. in height, should be used. This somewhat elaborate specification of the iron arc is necessary to ensure freedom from pole effect; (ii) a completely revised table of secondary standards, based almost entirely on the work mentioned above, was given. This is now known as *the 1928 scale*. It covers the same range as the 1922 scale but contains far more interferometer wave-lengths. The 1928 wave-lengths are all smaller than those of the 1922 scale, the amount ranging from 0·001 A. in the ultra-violet to 0·009 A. in the deep red. The differences are almost entirely accounted for by the presence of pole effect in the sources used by the earlier observers.

During the same period three important investigations on the solar spectrum were carried out with the object of establishing a new system of solar standards. One of these, which was carried out jointly by the Bureau of Standards and the Allegheny Observatory, gives the wave-lengths of seven hundred solar and atmospheric lines in the spectrum of integrated sunlight (i.e. the mixed light from the whole of the solar disc) over the range 3592 A. to 7142 A. The measurements were made with interferometers against the neon standards.

The two other sets of observations were made independently by St John and by Babcock at Mount Wilson. St John used the light from the centre of a solar disc 42 cm. in diameter, and carried out his measurements entirely with a plane-grating spectrograph of high dispersion, by comparison with the iron-arc standards. The region covered was 3592 A. to 6495 A.

Babcock made three series of observations, all with the interferometer. The first of these covers the region 4500 A. to 6900 A., and was made by comparing the light from the centre of a solar disc 17 cm. in diameter with that of a standard iron arc. The second, which covers the range 3710 A. to 4500 A., differs from this only in the use of integrated sunlight instead of light from the centre of the disc. In these two series there are about nine hundred lines. The third series gives the wave-lengths of five hundred lines in the infra-red (6900 A. to 9000 A.) on the neon scale.

These investigations enabled the Wave-lengths Commission of the International Astronomical Union to publish a list of three hundred and fifty-one recommended standards covering the range 3592 A. to 7122 A.

It must be admitted that the material from which the 1928 solar standards were derived is to a certain extent lacking in homogeneity, since some of the work was done on integrated sunlight and some on the light from the centre of the sun's disc. In the case of the atmospheric lines of the solar spectrum, there is of course no difference between the wave-lengths, but for the solar lines the wave-lengths of integrated sunlight are slightly greater on account of the limb effect. However, the difference is very small, about 0·001 A., except in the red; and even in the red, where it amounts to about twice this value, it is of no serious consequence because the breadth of the solar lines in this part of the spectrum makes it practically impossible to attain an accuracy better than about $\pm 0{\cdot}003$ A. in their measurement.

It should be mentioned that in 1928 the Mount Wilson revision of Rowland's *Preliminary Table of Solar Wave-lengths* made its appearance. This contains revised wave-lengths for all the lines given in Rowland's table on the 1922 scale. It is rather unfortunate that the 1928 scale was not employed in this exceedingly important work.

§ 9. MORE RECENT PROGRESS IN WAVE-LENGTH STANDARDS

The primary standard. The wave-length of the red line of cadmium has been determined interferometrically, by comparison with the metre, four times. The first comparison was made by Michelson in 1893, the second by Buisson, Fabry and Perot in 1906, the third by Watanabe and Imaizumi in 1928, and the fourth by

Sears and Barrell in 1932. These determinations are all in agreement to about 1 milliangstrom, and it seems that by far the largest source of error lies in the metrological process of comparing a line standard with an end standard, which is of course a necessity on account of the form of the primary standard. Since it is possible to make interferometric comparisons of wave-lengths with an accuracy ten times greater than that attainable in the absolute measurement of the wave-length of a spectrum line in terms of a material standard, it is essential to retain a spectrum line as primary standard.

At present the red line of cadmium is adopted as standard both by the International Astronomical Union and by the International Conference on Weights and Measures, but there is an important difference between the definitions of the source recommended by the two bodies for producing the line. While the International Conference on Weights and Measures adopted a somewhat stringent definition of the source (which practically amounts to stating that the condition used by Michelson must be reproduced as closely as possible), the International Astronomical Union stated in 1925 that any source which gave measurable fringes with retardations of 200,000 waves might be used. Some of the members of the International Astronomical Union Wave-lengths Commission advocated at the 1932 meeting that this definition should be abandoned in favour of the stricter one, but no resolution was passed on the matter.

Iron standards. During the last five years perhaps the most important contribution to our knowledge of the iron standards is the work of Burns and Walters[3,4] on the wave-lengths of the vacuum iron-arc spectrum. In two papers they give the wave-lengths of a great number of iron lines determined by means of the interferometer. The first paper gives the wave-lengths of about seven hundred lines, covering the range 2761 A. to 8800 A., determined by interferometric comparison with the neon standards. The average accuracy is about 1 part in 5,000,000. The second paper contains interpolated (interferometric) wave-lengths of about eight hundred lines. The source used for this work was a short iron arc *in vacuo*, operated at 6 to 8 amperes in order to photograph the weak lines in a reasonable time. The average accuracy of these wave-lengths is not so high as that of the lines given in the first paper. The range of the observations is 2150 to 3800 A. Between 2800 and 3800 A. iron lines which had been directly compared with neon (first paper) were used as standards. From 2150 to 2800 A., computed values of iron wave-lengths were used as standards, and also some copper impurity lines. When computed lines were used as standards their wave-lengths were increased by 0·001 A. to allow for the discrepancy between the observed and computed wave-lengths found in the first paper. The validity of this procedure is supported by the wave-lengths of the vacuum iron arc found by Kiess[5], but Jackson's[6,7] wave-lengths for the arc in air indicate that the computed values are the more nearly correct.

Some progress in the extension of the range of the iron arc in air was made by Jackson[6,7] who also worked with the Fabry-Perot interferometer and gave in two papers the wave-lengths of about 130 lines covering the range 2300 to 3500 A., with an accuracy of $\pm 0{\cdot}001$ A. The iron arc, which was of the type recommended by the

International Astronomical Union in 1928, was maintained with a current of $3\frac{1}{2}$ to 4 amperes, and varied in length between $1\frac{1}{2}$ and 2 cm. Only the light from the central 1 mm. was used for producing the lines. The wave-lengths found in this investigation showed that the earlier iron standards in the ultra-violet suffered from serious errors, in many cases greater than 0·01 A., and it is considered desirable for more measurements to be made as soon as possible.

Burns has shown that iron lines observed as impurities have slightly smaller wave-lengths than lines in the iron arc, but the difference is less than 0·001 A.

Krypton. The first spectrum of krypton is dominated by a group of very strong lines in the region 4273 to 4500 A. These are all extremely sharp and are situated in a very convenient part of the spectrum, and therefore appear to be eminently suitable as secondary standards. For the brightest ten of these lines there are in existence three independent wave-length determinations, which are in very good agreement. For this reason the lines were adopted as secondary standards by the International Astronomical Union in 1932. The wave-lengths are given below, as they have the interesting distinction of being the only eight-figure secondary standards recommended by the International Astronomical Union.

λ Recommended	Jackson[8]	Humphreys[9]	Meggers[10]
4273·9702	·9702	·9705	·9696
4282·9688	·9689	·9686	(·967)
4318·5522	·5522	·5522	·552
4319·5800	·5801	·5798	·580
4362·6425	·6425	·6429	·6422
4376·1220	·1221	·1217	·122
4399·9674	·9673	·9675	(·969)
4453·9179	·9179	·9183	·9174
4463·6903	·6906	·6897	·690
4502·3546	·3548	·3546	·354

Neon. The most recent work on the red lines of neon is that of Jackson[11]. He found that with high resolving powers, and interference paths up to 3 or 4 cm., his determinations were in almost perfect agreement with the values found at the Bureau of Standards. Both sets of measurements are given to eight figures and the agreement for nearly all the lines is within 0·1 or 0·2 milliangstrom. If, however, the resolving power is below about a quarter of a million, the apparent wave-lengths of all the neon lines become shorter, till for very low resolving powers (below about 100,000) they are about 2 milliangstroms shorter, after which there is no further change. This change in the apparent wave-lengths is caused by the blending of the satellite of Ne^{22} with the main line given by Ne^{20}. Jackson gives a table of corrections which may be applied for using the neon lines as standards with interferometers of medium and low resolving power.

Interferometric measurements in other spectra. (i) *Titanium.* There are three sets of observations on the spectrum of the titanium vacuum arc. These are by Brown[12], who measured more than a hundred lines between 4250 A. and 6261 A. against the primary standard; by Kiess[13], who measured more than three hundred lines in the region 2941 to 6743 A. by comparison with neon standards; and by Burns[14], who

gives wave-lengths of both the vacuum arc and the air arc in the regions 4055 to 4482 A. and 5739 to 6556 A. (ii) *Copper*. Burns and Walters[15] have made an extensive system of observations of the copper vacuum arc covering the range 2104 to 8092 A., and have given a term-classification for every line in their list of wave-lengths. The accuracy is about 1 part in 10^6 for the region below 2300 A., but it becomes greater with increase in wave-length. Kiess[13] has also determined the vacuum-arc wave-lengths of 26 copper lines, covering the range 2824 to 5782 A. These agree well with the wave-lengths given by Burns and Walters. (iii) *Calcium*. Burns[14] has measured twenty-five lines between 5265 and 8662 A., and Kiess[13], fifteen lines between 3833 and 6499 A. The agreement between lines common to both is quite good. Jackson[16] has made a special investigation of the wave-lengths of the H and K lines in the vacuum arc. (iv) *Solar spectrum*. At the 1932 meeting of the International Astronomical Union an important contribution to our knowledge of solar standards in the infra-red was made by Babcock and Hoge[17], who gave a list of two hundred and fifty interferometer wave-lengths covering the range 7333·684 to 9898·999 A. and seventy-five grating wave-lengths from 9908·29 to 11204·87 A. The grating wave-lengths were made against standards in a higher order and were given only to six figures. All the wave-lengths were on the 1928 scale.

§ 10. STELLAR RESEARCH

Absorption lines. The application of photographic spectrophotometry to astrophysical problems, though still in its infancy, has already resulted in conspicuous advances in our knowledge of absorption in stellar atmospheres. Recently the improved technique available has provided access to an entirely new field of knowledge —that concerned with the axial rotation of stars. That such rotation might give rise to Doppler broadening in stellar lines has, of course, long been suspected; that it actually does so in certain stars is known from the spectra of eclipsing binaries, in which an asymmetrical broadening occurs when one limb of the primary star is hidden near the time of eclipse.

The pioneer observations of Struve and Elvey at the Yerkes Observatory have suggested that equatorial rotation velocities of the order of hundreds of kilometres per second may not be uncommon. Unfortunately, however, their procedure is open to objection, as indeed they pointed out at the time. In fitting an observed line profile to one obtained by calculating the effect of a plausible rotation velocity on the original contour, they were forced to make two basic assumptions—first, that the original form of the line is known, and secondly that it is indeed rotation which distorts that form into the one observed. All necessity for such assumptions has now been removed by Carroll, who has shown that if rotation is present certain Fourier components will be absent from the mathematical formulae representing the observed profile. This effect is used to disentangle the rotation factor from others affecting the shape of the line, and to measure its magnitude. The method is shown to be strikingly free from errors due to instrumental and other distortions, and the results of its application to the lines He I 4387, He I 4472, and Mg II 4481 in the

stars 44 ι Orionis, 44 ζ Persei, and 39 λ Orionis give results in good agreement with those of Struve and Elvey. The critical test on the eclipsing variable Algol yields a result in remarkably close agreement both with the direct rotation velocity, determined as indicated above, and with the empirical value of Struve and Elvey. Evidently only the fringe of this fascinating subject has yet been touched upon.

An important statistical study of rotational broadening in B, A, and F stars has now been completed by Westgate, who uses the Struve-Elvey method to transform line-widths into equatorial velocities. By assuming an arbitrary distribution of stellar rotational speeds, a theoretical frequency curve for each spectrum class can be constructed, due regard being paid to the probable axial inclinations, and can be fitted to the observed-distribution graph. Stars of class A have, on the average, greater speeds of rotation than those of either class B or class F, appreciable rotation disappearing in the middle F stars. Little correlation is found between rotational velocity and absolute magnitude. Shajn has pointed out that the high rotation-speeds observed suggest that many single early-type stars are now in a state bordering on instability, whilst many early spectroscopic binaries may have passed through such a state relatively recently.

Bright-line stars. The study of bright-line stars, particularly those of type B, has of late attracted much attention from both the observational and the theoretical standpoints. The data accumulated during the past few years on stars of classes O, Be, and Ae have been collected by Merrill and Burwell, who add a statistical analysis of the distribution of such stars in spectral type, magnitude, and position. The tendency to emission seems greatest in B2 stars and then diminishes rapidly from B3 to A0. Detailed investigations of individual bright-line B stars suggest that many such stars are spectrum variables in which each hydrogen line consists of a central absorption line separating two emission components whose relative intensities are subject to cyclic variation. McLaughlin has advanced a tentative hypothesis of a rapidly rotating star, whose temperature varies cyclically, radiating into a slowly rotating atmosphere which is caused to pulsate by the varying radiation pressure. The spectrum variations are thus explicable purely as Doppler effects.

Merrill has discovered three peculiar stellar spectra in which bright lines due to He II, N III, and Fe II are superposed on TiO absorption bands. It is suggested that this apparently anomalous coexistence of high and low excitation spectra may result from extremely low pressure in the stellar atmospheres concerned.

Wolf-Rayet spectra. In no branch of stellar spectroscopy have greater strides been made in recent years than in the investigation of Wolf-Rayet spectra. Most of the previously unidentified lines are now known to be due to highly ionized carbon, oxygen, and nitrogen. Any one spectrum may contain lines arising from an atom in as many as four successive stages of ionization, the corresponding range of excitation being perhaps 100 V. A discovery which may be helpful in deciding the probable evolution of Wolf-Rayet stars is that two distinct parallel sequences exist, one exhibiting lines due to C III, C IV, O III, O IV, and O V but none due to nitrogen, and the other showing nitrogen in various stages of ionization but containing no evidence of carbon or oxygen. H, He I, and He II appear in both. Spectrophotometric

measures of total intensities have made possible the establishment of lower limits for the temperatures of Wolf-Rayet stars, on the assumption that their emission spectra are produced in accordance with Zanstra's theory of nebular luminosity. The values found are so high that they suggest a physical discontinuity between these stars and those of type B: the carbon sequence ranges from 61,000° to 103,000°, and the nitrogen sequence from 65,000° to 97,000°.

Forbidden lines. In a survey of metastable atomic states in many of the lighter elements, Boyce, Menzel, and Payne have compiled a list of forbidden lines, i.e. of inward transitions from such states, and compared it with lists of unidentified lines in the spectra of nebulae and of novae. Though many of the identifications are based on extrapolations and should therefore be regarded as tentative, others are almost certainly real. Perhaps the most interesting assignments are of the nebular lines λλ 3342, 3868, 3969 to Ne III, and λ 4724 to Ne IV. The possibility of a wide cosmic distribution of the supposedly rare gas neon is suggested by these discoveries, and a very recent attempt to identify lines of Ne I and Ne II in absorption in later-type stars has been successful, the intensities found suggesting a stellar abundance comparable with that of oxygen.

Novae. Amongst recent novae the most interesting newcomer is RS Ophiuchi, whose behaviour is described later. This star resembles T Pyxidis and a few others in showing repeated outbursts of radiation, and it is possible that a class of recurrent novae must be recognized. Such novae would not be explicable on Milne's theory, which attributes the appearance of a nova to a sudden release of energy following the occurrence of dynamical instability at a certain stage of stellar evolution, unless it were considerably modified.

Interstellar matter. The uniform distribution of interstellar matter required by Eddington's theory of diffuse atomic clouds has been definitely established by the analysis of the large amount of observational material accumulated at the Dominion Astrophysical Observatory, Victoria, B.C.: the relation between the distance of a star and the intensity of the interstellar lines appearing in its spectrum is exactly linear. The chief objection to the theory—that interstellar lines fail to appear in spectra of types later than B5—has been partially removed by the discovery that the binary H.D. 698 (type B9) shows both stellar and interstellar K lines in certain phases of its orbit. It seems likely that the presence of detached lines will be established in late-type spectra by examining stars at distances greater than 10,000 parsecs and having radial velocities great enough to shift the stellar lines well away from the interstellar ones.

The intensity of the sodium interstellar lines relative to those due to ionized calcium was calculated by Eddington as very low for any sodium abundance which is at all likely. Recent spectrophotometric measurements by Beals on the stars ξ Persei and 9 Camelopardalis, however, confirm visual observations in giving the ratios $K : D_1 : D_2 = 1 : 1 : 1$ to within 10 per cent. Fortunately the difficulty is not insuperable, and the presence of diffuse matter, uniformly distributed throughout the galaxy with a density of 10^{-25} gm./cm.^{3} and a temperature of 15,000° K., may be taken as established.

§ 11. SOLAR RESEARCH

As the only stellar body from which we can obtain detailed information, the sun naturally presents a fruitful field for research.

Chromosphere. The study of the distribution of elements in the solar atmosphere has been carried a step farther by Mitchell and Williams, who have developed an entirely new method of probing the chromosphere spectroscopically. In objective prism flash spectra, as is well known, the tips of lines of different lengths show the various heights to which the corresponding atoms extend in the chromosphere. By an ingenious combination of the multiplet-intensity formulae with a knowledge of these heights for all the lines in a given multiplet, it has been possible to derive, for several elements, the density-distribution in the lower chromosphere. The density-gradients found agree well with previous results, and are of the same order for all the elements studied, though these include atoms as different from one another as H, Fe, and Ti^+. This indication of an almost perfect mixture of the solar gases suggests that Milne's theory of chromospheric support by selective radiation pressure can be only partially true, turbulence probably playing a greater part than has hitherto been realized.

Corona. No account of recent solar research would be complete without a reference to the work of Lyot on the corona. By taking extraordinary precautions to reduce scattered light, both by careful design of his telescope and by observing at an elevation of nearly 10,000 ft., he has been able for the first time to photograph the prominences and corona in full sunlight without a spectrograph. Spectrograms of the corona show the red and green lines plainly, whilst the dispersion (1·2 A./mm.) at his disposal enabled their wave-lengths to be measured to a few hundredths of an angstrom.

The rejection of "nebulium" in favour of ionized oxygen and nitrogen as the source of the mysterious lines occurring in the spectra of galactic nebulae has strengthened the conviction that the hypothetical "coronium" is also a familiar element in an unfamiliar guise. The coronal lines λλ 6776, 6704, 6374, and 5303 have recently been assigned by de Bruin and Menzel independently to the atom of neutral oxygen; but the evidence is by no means conclusive, and weighty objections to the proposed classifications have been put forward. It seems that the origin of the lines is as much a mystery to-day as it was when they were discovered 65 years ago.

The discovery of five coronal lines, λλ 3987, 4086, 4231, 5303, and 6374, in the spectrum of the variable star RS Ophiuchi during its 1933 outburst promises to be of the first importance for investigators of the solar corona. The nebular lines typical of the spectra of novae reached their highest intensity about 3 weeks after maximum in this star, and during their decline the coronal lines appeared gradually, their presence being definitely established 7 weeks after maximum. This is the first time that coronal lines have been found in a source other than the corona itself, and the discovery appears likely to help considerably both in tracing the origin of the lines and in deciding the physical conditions concerned in their production.

Absorption lines. Until recent years little attention was given by spectroscopists to the intensities of spectrum lines, visual estimates being considered sufficient for most purposes. It is now, however, becoming increasingly evident that, in astrophysical work particularly, accurate intensity-measurements are capable of yielding results of the first importance. In the next few years determinations of intensities of spectrum lines may come to rival in significance measurements of their wave-lengths. At present, the astrophysical interest in such determinations lies in two directions: first, variations in the relative intensities of lines in connected groups or multiplets may be expected to throw light upon differences of excitation in different sources; and secondly, since the contour of a line, i.e. the distribution of intensity across its width, must be intimately connected with the mechanism of the absorbing atmosphere in which the line originates, contour-measurements will probably prove of vital importance in the study of stellar evolution. It is not surprising, then, that with the recent advances of technique in methods of photographic photometry has come theoretical progress requiring for its full development still greater instrumental improvements.

A new method of obtaining line profiles, due to Shane, involves the use of a Fabry-Perot etalon crossing a grating spectrograph. The increased resolving power thus employed, whilst offering little advantage in the determination of equivalent breadths, increases the accuracy obtainable in measuring central intensities.

An exhaustive microphotometric calibration of the Rowland scale of solar intensities has now been almost completed at Utrecht. A preliminary survey shows that the equivalent breadth corresponding to a given Rowland intensity does not remain constant at all wave-lengths: for all intensities, a general rise from λ 4000 to λ 5500 is terminated by a sharp drop of 70 per cent in the region $\lambda\lambda$ 5500–5900, beyond which the steady rise is resumed.

The microphotometry of the solar spectrum from λ 4040 to λ 4390 has also been carried out by Woolley. Equivalent breadths and central intensities for nearly 500 lines are given, and it is claimed that the data are free from the instrumental errors present in previous determinations of line contours. The investigation establishes beyond question the existence, hitherto regarded as doubtful or even denied, of deviations in solar multiplets from the normal intensity-summation rules. Thackeray's intensity-measurements in an iron multiplet in the green confirm a similar effect in stellar spectra: a marked weakening of strong components relative to weak ones is established in K5 and M stars. Both authors incline to an explanation involving interlocking between different lines in a multiplet. Briefly, this hypothesis suggests that at high levels in a stellar atmosphere the continuous photospheric radiation, deficient as it must be in wave-lengths corresponding to the strong lines of a multiplet, can bring about fewer of the corresponding upward transitions amongst the absorbing atoms than theory would indicate. The radiation deficiency being less for weak lines, the efficacy of the continuum in causing these transitions is diminished to a smaller degree, and consequently the weak lines appear enhanced.

The problem of the high central intensities observed in Fraunhofer lines remains

unsolved. Any theory assuming merely scattering and absorption in the solar atmosphere requires complete blackness at the centre of a line, whilst even Eddington's superelastic collisions do not fully account for the facts. An important suggestion by Evershed which seems likely to throw some light on the matter attributes the unexpected central brightness to emission from hot jets of gas in a turbulent chromosphere. In other words, the conflict between theory and observation is explained as due to a departure from the theoretical condition of horizontal stratification in the solar atmosphere—a departure, moreover, which is independently confirmed from other lines of evidence.

Rotation. It has long been recognized that marked discrepancies are present in different observations of the solar rotation by spectroscopic methods; the latest addition to the obscurity is contained in the announcement that the 1932 Pulkova results fail to confirm the increase in the equatorial velocity recently noticed at certain other observatories. Local horizontal currents in the solar atmosphere might explain such discrepancies, and have indeed been postulated for that purpose. Objective evidence of their existence has now been obtained by Evershed, who finds irregularities in Fraunhofer wave-lengths which correspond to random horizontal velocities of the order of 0·4 km./sec. This value is surprisingly high, being 20 per cent of the total equatorial velocity, and is sufficient to reconcile the discordant results. The moral appears to be that observations must be made more frequently in order that the effect of the disturbances may disappear in the mean.

§ 12. INSTRUMENTAL TECHNIQUE IN ASTROPHYSICAL SPECTROSCOPY

Two major instrumental advances made during the past year call for mention; they are concerned with extending the range of astronomical observations towards longer and shorter wave-lengths respectively. The new photographic plates, which make the former extension possible, have already been referred to.

Infra-red photography. The region λλ 8320–8900 is free from general terrestrial absorption, and plates with maximum sensitivity here have been used with conspicuous success in astrophysical spectroscopy. The most prominent features of stellar spectra revealed by the new plates include lines attributed to hydrogen, Paschen series, in class A stars; Ca II–8498, 8542, 8662 in A to M stars; O I–7771, 7774, 7775, 8446 in A to F stars; Ca I, Ti I, Fe I in K to M stars; and bands due to CN in N stars and to TiO in M stars. Investigations of bands in planetary spectra in this region have established the presence of carbon dioxide in the atmosphere of Venus, and of ammonia and methane in both Jupiter and Saturn. The atmosphere of Mars, if it contains oxygen at all, has been shown to do so in amounts which are negligible compared with the terrestrial abundance. The near future will undoubtedly see the extension of the present limit beyond the next region of atmospheric transparency (λλ 9800–11,000), possibly by the use of xenocyanin plates, and there seems no reason why photographic research should stop even there.

Aluminized mirrors. It is well known that the transmission band of metallic

silver at λ 3200 is responsible for the poor performance in this region of reflecting telescopes employing silvered surfaces. The technique of coating glass surfaces with a suitable film of aluminium has now been perfected, and the operation has already been successfully performed on the 36-inch Crossley mirror of the Lick Observatory. The metal is evaporated on to the glass surface in an atmosphere at a pressure of 10^{-4} mm. The resulting layer is 0·1 μ thick and can be deposited with extreme uniformity. Its reflecting power is scarcely inferior to that of silver in the visible region, whilst at λ 3250 it is seven times, and at λ 3000 ten times, as great. Aluminized mirrors, therefore, will enable the solar spectroscopist to photograph with ease to the limit set by terrestrial absorption at λ 2950, whilst the speed of direct photography of faint stellar objects will obviously be improved considerably. The subsidiary properties of the aluminium film appear to be very good: it does not tarnish, it is tenacious enough to be cleaned in safety with soap and water, and it does not scatter light.

REFERENCES

(1) Mack, Stehn and Edlén. *J. opt. Soc. Amer.* **22**, 245 (1932).
(2) Edlén. *Nova Acta Soc. Sci. upsal.* IV, **9**, No. 6.
(3) Burns and Walters. *Publ. Allegheny Obs.* **6**, No. 11.
(4) —— —— *Publ. Allegheny Obs.* **8**, No. 4 (1930).
(5) Kiess. *Bur. Stand. J. Res.*, Wash., **1**, 75–90 (1928).
(6) Jackson. *Proc. roy. Soc.* A, **130**, 395 (1931).
(7) —— *Proc. roy. Soc.* A, **133**, 553 (1931).
(8) —— *Proc. roy. Soc.* A, **138**, 147 (1932).
(9) Humphreys. *Bur. Stand. J. Res.*, Wash., **5**, 104 (1930).
(10) Meggers. *Sci. Pap. U.S. Bur. Stand.* **17**, 193 (1921).
(11) Jackson. *Proc. roy. Soc.* A, **143**, 124 (1933).
(12) Brown. *Ap. J.* **56**, 53 (1922).
(13) Kiess. *Bur. Stand. J. Res.*, Wash., R.P. No. 4 (1928).
(14) Burns. *Publ. Allegheny Obs.* **8**, No. 1 (1930).
(15) Burns and Walters. *Publ. Allegheny Obs.* **8**, No. 3 (1930) and Supplement (1931).
(16) Jackson. *Mon. Not. R. astr. Soc.* **93**, 98 (1932).
(17) Babcock and Hoge. *Trans. Int. astr. Un.* **4**, 82 (1932).

OPTICS

By MEMBERS OF THE STAFF OF THE NATIONAL PHYSICAL LABORATORY

§ 1. Optical instruments, by T. Smith, M.A., F.Inst.P., F.R.S.
§ 2. Refractometry, by J. S. Anderson, M.A., D.Sc., Ph.D., F.Inst.P.
§ 3. Vision, by J. Guild, A.R.C.S., D.I.C., F.R.A.S., F.Inst.P.
§ 4. Colorimetry, by J. Guild, A.R.C.S., D.I.C., F.R.A.S., F.Inst.P.
§ 5. Photometry, by H. F. Buckley, M.Sc., F.Inst.P.

---

§ 1. OPTICAL INSTRUMENTS

Comparatively few physicists have carried the study of optical instruments beyond the elementary stage needed to enable them to use the more familiar instruments intelligently. Their general unfamiliarity with the art of designing and constructing these physical tools makes it desirable to preface an account of recent progress in this field with an introduction which begins much further back than is necessary in most branches of physics, including other branches of optics.

If we contrast optical knowledge and practice one hundred years ago with that at the end of the nineteenth century we can hardly fail to be conscious of many changes. Among the more noticeable are the greater variety of types of instrument, the vast improvement in defining power, in light-grasping capacity, and in the extent of the fields they cover. Some of these improvements are primarily due to advances in the manufacture of glass for optical purposes; this applies particularly for example to the larger telescope objectives. Others have come about as an indirect result of new scientific knowledge; photographic lenses with moderately large apertures and fields of great angular extent furnish obvious examples. Still other improvements have resulted from more exact knowledge of optical theory, microscope objectives of large numerical aperture being outstanding illustrations. Remarkable as these changes are, they would perhaps have been less surprising to those learned in optical matters some seventy years earlier than a change of a different kind. There is ample evidence that at the beginning of this period the master craftsmen were generally without any knowledge of the theoretical principles governing the construction of the instruments they made. Such theoretical investigations as were required were carried out by leading natural philosophers, who recorded their results in a form which could readily be utilized by the artists in the workshop. By the end of the century all this had changed. Manufacture was then, as now, usually preceded by extensive calculations of a laborious character, carried out by members of the optician's staff. The calculations on a new lens may take from six to twelve months, and sometimes considerably more. Under these circumstances it is not surprising

to find the plaint early in this century that academic optics was definitely behind the optics of the best instrument-makers.

If a manufacturing optician had been asked in 1900 to name the workers who had contributed most to the progress of optics in the preceding seventy years it is not unlikely that he would have mentioned Gauss, von Seidel, and Abbe. All three contributed to the theory of optics, but only Abbe applied the theory to the construction of instruments. If we inquired further in what way these three had helped the practical optician, we should learn that Gauss had systematized the elementary (i.e. non-aberrational) theory, with the result that any non-telescopic lens system could be considered as a single lens. On further inquiry we might reach the conclusion that this work in fact affected the optician to a less extent than he had supposed. The contribution of von Seidel lay in a systematic exposition of the first order aberrations and the construction of expressions by which these could be evaluated for an instrument of which the constructional particulars were already known. We should probably also be informed that this work revealed an aberration previously unknown, but since described as coma. Close examination would presumably show that the optician—at least if a British optician—did not make any use of Seidel's expressions, but, knowing the nature of the aberrations, determined their magnitude in an entirely different way. The methods used, we should find, were either the experimental examination of model instruments made up in the workshop, or the mathematical analysis of a hypothetical design on paper on the basis of several rays traced trigonometrically through the system. In either event progress was made by trial and error, tentative modifications in the design being based on the previous experience of the investigator, or, perhaps less frequently, on expressions such as those given by Coddington before our period opened. These expressions would be interpreted as qualitatively reliable, but only very roughly correct for actual magnitudes. Thus the real influence of von Seidel's work was probably not great.

Coming now to Abbe we find contributions of immediate and undeniable value to the designer. The trigonometrical method of designing is extremely laborious, and Abbe's sine theorem

$$\operatorname*{Lim}_{y\to 0}\frac{y'}{y}=-\frac{\mu\sin\psi}{\mu'\sin\psi'},$$

which enables the size of the image of a very small object on the axis of an instrument produced by a narrow annular zone of the lens to be determined from the initial and final directions ψ and ψ' of a single ray traced through that zone, greatly reduced the amount of calculation which was necessary. Especially was this theorem of value in the design of large aperture microscope objectives. In addition Abbe expounded a particular way of regarding imagery in a microscope: we shall have occasion to refer to this shortly.

We may then summarize by saying that the only important addition made to optical theory as appreciated by our optician was due to Abbe. We now know that what was not in Coddington was fully described by Hamilton (though it should be said that Hamilton's writings would certainly be too difficult for most optical readers), and that our optician was incorrect in supposing von Seidel to have been

the first investigator to treat of coma. We also note that the designer of this time was thinking solely in terms of rays, and judging of the quality of his instrument from the size of the area in which they intersected the assumed image plane. Rayleigh's work, with its emphasis on waves and wave-lengths, seems to have gone unnoticed by technical workers; at least there is no evidence that it was appreciated.

The one instrument in which some notice was generally taken of the wave character of light was the microscope. This can be traced largely to Abbe's influence. Abbe observed a simple connection between the resolution of the image of an illuminated grating and the location of the interference images of the source produced by the grating with respect to the aperture of the microscope objective. With this as a starting point he described a way of regarding image formation in a microscope which enables practical microscopists, without studying the wave theory (by which the validity of all such conclusions must in the last resort be established), to recognize in many cases whether or not the conditions for the resolution of detail were or were not satisfied. That Abbe's teaching accorded with the wave theory was shown in special examples by Lord Rayleigh. The effect on some of Abbe's disciples of this teaching, which was often mistakenly regarded as an independent and sufficient theory of image formation for the microscope, was curious. It came to be seriously held that images formed by the microscope were different in nature from those formed by a telescope or a photographic lens. Apparently light could recognize that the instrument it had entered was a microscope, and thereupon manifested the characteristics of waves and formed an image by interference. If on the other hand it concluded that it was traversing an instrument of a different kind, it behaved like a corpuscle and the image corresponded to the geometrically determined intercept of rays on a plane. The dilemma is, superficially at all events, similar to one that has invaded physics in recent years. Though it is unlikely that any competent optician could be found to-day to uphold this view, its effects are still with us. The mere fact that outstanding aberrations are still commonly expressed in terms of geometrical intercepts, and their seriousness judged by such magnitudes, shows what a hold the corpuscular rather than the wave notion of light still has on lens designers and technical writers. Authors of text-books on light and optics show the same bias when they propose the diameter of the "least circle of confusion" as a measure of the seriousness of spherical aberration in instruments.

At the beginning of the present century then the superiority of the instrument-makers' optics depended on the accumulation and application of information derived from experiment, or more frequently from accurate ray tracing, rather than on any systematic extensions of the general theory of optical instruments. Since then much progress has been made in many directions, of which only a few can be mentioned here. First in importance is the amazing advance which has been made in the performance of commercial instruments. To take examples, microscope objectives of all powers, the highest of which represented the summit of optical achievement in the last century, have unquestionably been improved. Recently prismatic binoculars have been made which transmit far more light over the whole field than earlier designs, and this without any loss in defining power. The require-

ments of cinematography have led to the development of photographic lenses which will take pictures able to stand the great enlargement required, with apertures of diameters which may be as great or even greater than the focal length of the lens itself. The magnitude of these achievements can hardly be appreciated by those without some detailed knowledge of the difficulties surmounted. Improved materials, tools, and methods of manufacture have played a part in this advance. The glass manufacturers have assisted by making glasses in considerable variety and of a transparency and homogeneity unattained before: they have also co-operated by producing glasses with assigned optical constants for special purposes. But most of all the improvements rest upon improved designs ultimately reached by very much the same kind of trigonometrical work as was in use thirty years earlier.

A good idea of much of this work may be gained from Conrady's[1] book on *Applied Optics*. This book also illustrates well the extent to which physical rather than geometrical concepts have been employed in recent years in the computation of lenses. The concept of stationary time and the significance of the wave-length play a subordinate, though by no means negligible, part in the work. The main emphasis is still on the tracing of rays, and the calculation of path-difference enters as an afterthought to the main computation. The remarkably fine achievements of the optical industry to-day then owe less than might have been expected to improvements in methods of calculation and to comparatively new theoretical developments.

Nevertheless the additions to our theoretical knowledge which have been made of late are not small either in number or in importance, and they are likely to prove of increasing significance in future instruments. Except in Germany they have come mainly from workers outside the industry rather than from those within it. Among the reasons for this change we may reasonably assign some weight to the fact that several have arisen from, and are most readily appreciated with the aid of, the wave-view of image formation; moreover they have been derived by the use of mathematical tools of a different kind from those used for everyday elementary purposes. It is of course possible to present the new knowledge in terms of the traditional theory, and this can often be done with considerable ease. The *Report*[2] *on the Teaching of Geometrical Optics* recently issued by the Physical Society shows that the derivation of exact equations need not be in the least more complicated than that of the usual approximate relations, and that new discoveries, such as the generalization of the sine law for image fields of any size,

$$\mu y \sin \psi + \mu' y' \sin \psi' = yy' a,$$

where μ and μ' are the refractive indices of two media in which conjugate points are distant y and y' respectively from the axis, ψ and ψ' are angles with the axis by a ray through these points, and a is the power of the instrument for this ray, may also be established in elementary ways. Nevertheless the presentation of investigations in less familiar ways is not attributable to perversity on the part of the authors. Special algebras are used when they are particularly appropriate. An excellent example of the advantages gained by a wise choice is afforded by the recent translation[3] of Hamilton's papers on systems of rays into German. The mathematics, apart from

certain literal changes, has been reproduced in its original form: but the translation is accompanied by a pamphlet in which the same ground has been traversed employing a vector notation. The gain in clearness and conciseness is indeed remarkable. The algebra of matrices has been applied in recent years to many optical problems, and has been found particularly suitable for the computation of the properties of optical instruments. The changes involved in these departures from established methods are by no means slight, and great interests have been vested in trigonometrical work. It is therefore natural that these methods should be slow in their effect on industrial practice. The advice not to put new wine into old skins, which may here be applied in more ways than one, is not to be forgotten, but equally the need to bottle the new wine for use when it has matured must not be overlooked.

A general characteristic of much of the recent theoretical work is the adoption of Hamilton's characteristic function or one of the associated functions as a foundation. This method liberates us from all risk of overlooking any aberration in the way that coma was overlooked a hundred years ago, enables us to avoid introducing superfluous functions as coefficients, and permits us to deal with skew rays as readily as with coplanar rays, and with aspherical surfaces as easily as with spheres. In calculations it is often convenient to work with the second differential coefficients of the characteristic function as the constituents of a matrix rather than use the function directly. These developments have arrived very opportunely, since the expansion of apertures towards their limiting values has increased the importance of knowing how skew rays behave and of ascertaining how aspherical surfaces of revolution can best be utilized in these extreme constructions.

The extent to which photographic emulsions and lenses jointly have increased in rapidity will be appreciated from the fact that apparatus for taking 2500 photographs a second is a commercial product. The relative perfection of the two is not without interest. If we take the best modern lenses of apertures that would have been considered very large ten years ago we find that they will render definition finer than the photographic plate can utilize. There is therefore every reason why plate and film makers should endeavour to produce materials rendering finer detail. On the other hand there is some doubt whether the most rapid commercial lenses of to-day yield definition as fine as that with which the plates can deal. There is therefore ample room for further improvements in lens construction also.

Some reference may be made here to publications on optics other than papers communicated to scientific societies, which are almost invariably too technical to be discussed in a general review. Mention has already been made of the *Report on the Teaching of Geometrical Optics*[2]. The recommendations made in this report should, if adopted in schools and colleges, go far to remove the conflict of conventions which has hindered students in the past. In Germany much progress seems to have been made in a movement to standardize the symbols used in optics. The proposals may prove convenient for technical work on the established model, but their enforced application over a wider field would inevitably handicap rather than assist. The present writer would strongly deprecate the adoption of similar proposals in this country.

A long-standing defect in German optical literature has been repaired by the translation[3] of Hamilton's century-old papers on systems of rays. In view of the tendency to adopt the stationary path as the fundamental concept in modern work this translation should be widely used. Mention has already been made of the conversion of the mathematics into vector form. This is only one of a series of important monographs which have appeared recently in Germany, and which should be consulted by all interested in the growth of optical knowledge. Among these mention must be made of valuable contributions by H. Boegehold[4], M. Berek[5], M. Herzberger[6] and J. Picht[7], while on spectacles mention should be made of a recent publication by M. von Rohr and H. Boegehold[8]. An excellent text-book on general optics, which may be recommended to those able to read Dutch, has been written by Zwikker[9]. In the aggregate these books give a fair picture of present-day optical theory.

If we now turn from lens theory to means of testing lenses, we find a widespread tendency to adopt interference methods. Almost every important industrial country has one or more methods to its credit. The general impression derived from reading descriptions of these methods is that a wide choice is offered, with simple apparatus requiring much time, work and skill at one end of the scale, and at the other end costly apparatus which is quick and simple to use and to interpret. As far as can be learnt the Hilger camera-lens interferometer, which is a decidedly expensive piece of apparatus, is more widely used than its less expensive rivals.

The most interesting development in methods of testing optical instruments during the last year is Zernike's[10] phase-contrast test. This is essentially an improvement of the knife-edge test, the improvement being of a radical character. Substantially the defects of the knife-edge test are removed by altering the relative phases of an image and its first diffraction band by a quarter of a period. The equipment required is simple, and the test is applicable among others to instruments of so large a size that the Hilger method is out of the question. Details of the experimental technique have been given by Burch[11]. The theory of this test serves as a further illustration of the misleading character of results based on a purely geometrical treatment of optical instruments. The limitations of such results and of tests based upon them is by no means as widely recognized as it should be. To take an example, the present writer would not attach much weight to the figure of merit of an astronomical objective given by Hartmann's criterion.

Several papers have appeared lately dealing with the theory of curved gratings. It is worth noting that except in one particular, which from the point of view of the mathematical theory is almost trivial, there is no distinction between gratings and the ordinary reflecting or refracting instrument. The general theory of ordinary instruments can be adopted and applied to gratings, and the considerations which have been used in ascertaining what tolerances are permissible in the construction and in the use of lenses apply unaltered to gratings also.

In spite of the great amount of attention it has received the last word has not yet been written on optical images formed by the microscope. Important differences of opinion exist on the relation between the images of luminous and non-luminous

objects, and on the comparative resolution attainable with direct and with dark-ground illumination. Reference may be made to papers by Martin[12, 13], Berek[14] and Münster[15]. The present indications are that a relation useful to practical microscopists can be formulated between the resolution obtainable with luminous and non-luminous bodies, but that the dark-ground problem needs thorough experimental examination before any general conclusion can be reached. The theoretical work carried out hitherto on the latter problem is subject to severe limitations, and the evidence that the theory in its present form is applicable to the real problem consists only of a small number of measurements.

In the field of visual optics an interesting investigation on the form of the horopter has been carried out by Ames, Ogle and Gliddon[16].

§ 2. REFRACTOMETRY

The ancillary measurements of greatest importance in the production of optical instruments are those of refractive index. Measurements of refractive index are in fact used for many purposes, and the order of accuracy required in the measurements also varies greatly with the application. The Rayleigh interference refractometer for gases and the Abbe refractometer for liquids are well-known standard instruments and are applied in fields where high accuracy and comparatively low accuracy respectively are needed. We shall only deal in this article with measurements of the refractive index of glass, as it is only in this field that important changes have been made.

For a long period it had been the custom to measure the refractive indices of optical glass for the three hydrogen lines *C*, *F*, and *G* of wave-lengths 6563, 4861, 4341 angstroms and the *D* line of sodium of wave-length 5893. The normal requirement (for special purposes higher accuracy was necessary) was that measurements should be reliable absolutely to about one unit in the fourth decimal place, and differences of index for these lines reliable to about two or three units in the fifth decimal place. These lines were open to the objections that some were unsuitable for measurements of high accuracy and that their distribution in the visible spectrum was unfavourable for yielding accurate knowledge of the run of the dispersion. Within the last few years the principal glass-makers have changed their reference lines and given much more complete information. For example, the most recent Chance-Parsons list gives for about eighty types of glass refractive indices and differences for the following lines:

List designation	Element	Wave-length
b	He	7065
C	H	6563
d	He	5875
e	Hg	5461
F	H	4861
g	Hg	4359
h	Hg	4047

In addition, for the convenience of old users, values for the D and G' lines are added. The line denoted here by d replaces D as the general reference line, for which it is much more suitable, since the use of a double line like D is apt to cause inaccuracy in some measurements. In this list a very convenient system of reference numbers has also been adopted. Each number consists of six digits, the first three being the three leading decimal figures of μ_d, the index for line d, and the last three the leading digits in the constringence ν, defined by $\nu=(\mu_d-1)/(\mu_F-\mu_C)$. The number of a glass therefore gives the optician a reasonably accurate idea of the purposes for which it can be used in the construction of instruments.

For ordinary industrial needs measurements on glass may be made with a carefully calibrated and controlled Pulfrich refractometer. Only a small specimen is needed and its preparation is not expensive. If concordant results are to be obtained from different instruments, the necessity for special calibration should not be disregarded. Special tables are used in deriving the index of the specimen under test, and as the method involves refraction at grazing incidence it is important that

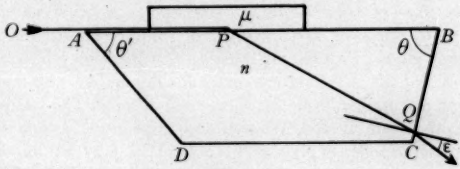


Figure 1.

these tables should be based on the exact properties of the glass used in the instrument. A manufacturer of refractometers will inevitably tend to make one set of computations serve for several instruments though the glass may be slightly variable. As trouble has frequently arisen through differences in the results obtained on the same specimen of glass with different instruments, and variation in the index of the glass used in the instrument itself is a likely cause of these differences, it is of interest to notice a recent development[17] in which the grazing refraction is retained but the need for tables is avoided. The principle involved is shown in figure 1. The glass to be tested of refractive index μ rests upon a standard block $ABCD$ of index n. The ray OP is refracted at grazing incidence and emerges at an angle ϵ with the normal to the face BC. With the angles as shown it is readily found that

$$\mu=\sin\theta\,\sqrt{(n^2-\sin^2\epsilon)}-\cos\theta\,\sin\epsilon.$$

With the usual design of Pulfrich refractometer $\theta=90°$, but there are advantages in adopting other angles. For instance an angle of 60° enables the index of the block to be measured on a spectrometer. In the new modification, rays are incident in turn in opposite directions, and the angles θ and θ' differ. The formula for the refractive index may then be put in the form

$$\mu=\frac{\sin\epsilon\,\sin\theta'-\sin\epsilon'\,\sin\theta}{\sin(\theta-\theta')}.$$

One suggestion is to make $\theta = 90°$, $\theta' = 45°$, so that we have

$$\mu = \sin\epsilon + \sqrt{2}.\sin(-\epsilon').$$

It will be observed that in this case ϵ' will always be negative on the convention we have adopted here. In an alternative construction $\theta = 90°$, $\theta' = 67° 30'$, a roof edge is provided along AD, and part of the face BC is silvered. Part of the beam refracted at P then emerges at Q at the angle ϵ, and part is reflected in turn at BC and the roof, finally emerging from the face BC at an angle ϵ'. The refractive index is then determined as before by the equation

$$\mu = \sin\epsilon + \sqrt{2}\sin\epsilon'.$$

Under the best conditions the Pulfrich refractometer is capable of yielding an accuracy of about 0·0001 in the refractive index and 0·00001 in the dispersions. When greater accuracy than this is required—for example when the index is required to about one unit in the sixth decimal place—an accurate spectrometer or other special apparatus is employed and the reliability of the result depends among other factors on the use of a sufficiently large prism with surfaces worked very accurately to the plane. It is desired to emphasize these conditions, as a series of papers by Tilton[18, 19, 20] published in the last three years leaves the impression that a small prism with surfaces decidedly curved will serve equally well if not better. Such a conclusion is justified neither by theory nor by experience. Some factors of much importance are neglected in these investigations, and assumptions which would not necessarily hold in practice are made without explicit acknowledgment. As far as can be judged from internal evidence the preference for small prisms is attributable to the assumption that only poorly corrected collimator and telescope objectives will be available for these measurements. In reality there should be no difficulty in obtaining lenses of suitably large aperture for this work which are sensibly perfect in their corrections. In the opinion of the writer the highest accuracy is only attainable if every care is taken to ensure that both tools and specimens are as perfect as possible, and to avoid every likely source of error in the course of the measurements. Tilton's papers contain an analysis of many of the errors that may arise.

§ 3. VISION

A useful starting point for any survey of recent work relating to the visual process is the *Discussion on Vision* held under the joint auspices of the Physical Society and the Optical Society in London, June 1932, and published as a separate report[21].

The report contains papers dealing with nearly all aspects of visual reception by authors of many nationalities. Physicists, physiologists and psychologists, prominently associated with the study of visual problems from these various standpoints, contribute their individual points of view and discuss those of the other contributors from a common platform.

It would be impossible, for reasons of space, to give adequate notice to each of the important papers in this compendium, and unfair to select a few of them.

Suffice it to say that anyone wishing to become conversant with the direction of present-day research into the processes of vision should study and digest this report from cover to cover. It is not suggested that the report, as a whole, provides either a complete picture or tells a consistent story. On the contrary, many fundamental divergences of view and disagreements on matters of fact are revealed in its pages; but the value of the report as a whole is many times greater than the sum of the values of the individual communications, important though many of these are, inasmuch as it enables the reader to bring into a single perspective the various aspects of the same problems which dominate the picture as seen from the separate viewpoints of the physicist, physiologist and psychologist, and to assess the extent to which confusion and misunderstanding have crept into the discussion of visual matters through the use by different groups of investigators of an essentially common vocabulary with an essentially different significance. If this report fulfilled no other purpose than revealing the different meanings which the same statement may sometimes have when made by, say, a physicist on the one hand and a psychologist on the other, and the impossibility of an intelligent interchange of knowledge between them until they use a common set of postulates and definitions, its publication would have been fully justified.

Since this report appeared, many important papers on various problems of visual reception have been published. Fresh information has been forthcoming on the sensitivity to qualitative changes in the stimulus, on visual acuity, and on the phenomena of intermittent vision, in addition to various miscellaneous matters.

A complete account of the determination of the sensitiveness of the eye to differences in stimulus purity* carried out at the Imperial College of Science by L. C. Martin, F. L. Warburton, and W. J. Morgan, was published as a Special Report of the Medical Research Council[22]. Determinations were made of the total number of just perceptible steps from white to the spectrum for a series of wave-lengths, and also of the variation of size of the purity limen, that is, the minimum change in purity which is perceptible, for different values of the purity. The values for the first step from white are found to be in close agreement with previous results of Priest and Brickwedde when corrected by a suitable factor to compensate for difference of field size. The curves showing the variation of the purity limen with purity exhibit interesting features, the most notable being that the purity limen for stimuli of the same hue wave-length is smallest in the neighbourhood of white and of the pure spectral colour, and has a maximum value for purities in the neighbourhood of 50 per cent.

D. B. Judd[23] has determined sensibility to colour temperature change as a function of temperature from 1800° to 11,000° K. and finds that $\delta T/T^2$ is practically constant.

E. P. T. Tyndall[24] has examined the chromatic sensibility to variation of hue

* *Stimulus purity* is the term used by these authors to denote the proportion, measured in luminosity units, in which monochromatic light of an appropriate wave-length must be mixed with a specified white light in order to match the colour of a given stimulus. Some authors term this ratio *colorimetric saturation*. The wave-length of the monochromatic constituent of the mixture is generally referred to as the *hue wave-length* or *dominant wave-length*.

wave-length as a function of purity. He found that for wave-lengths of 635, 580, 530 and 493 mμ. the sensibility decreases continuously with decrease of purity, but in the neighbourhood of 455 mμ. the results showed a secondary maximum at a purity of about 15 per cent. This was shown to be dependent on the spectral distribution of the heterogeneous stimulus.

A new determination of sensibility to change of hue wave-length throughout the spectrum has been made by W. D. Wright and F. H. G. Pitt[25]. The general features of the results agree with those obtained by Jones and Steindler, with the exception that no trace of a secondary maximum of sensibility near the red end of the spectrum was found.

Problems connected with adaptation have received attention both directly and indirectly. S. Hecht, G. Ward and C. Haig[26] find that the process of dark-adaptation takes place in two parts "sharply separable in time and characteristics." The first part takes about 3 minutes, and reduces the threshold to about $\frac{1}{\text{x}0}$ of its initial value. The second stage, which occupies about 30 minutes, reduces the threshold to about $\frac{1}{\text{x}0\,0\,0\,0}$ of its initial value. From a consideration of the effects observed in various retinal areas, it is concluded that the first stage is associated with the behaviour of the cones and the second with that of the rods.

W. S. Stiles and B. S. Crawford[27] have determined the liminal brightness increment as a function of wave-length for different conditions of foveal and parafoveal vision. The test spot consisted of a small rectangular strip of approximately monochromatic light within a small circular area 1·28° in angular diameter which was dark except for the test strip. The state of adaptation or "condition" of the retina was determined by the illumination of a large circular area 30° in diameter surrounding the central field. The intensity and colour of this "conditioning" stimulation could be varied at will. The threshold brightness at which the test strip becomes visible was determined throughout the spectrum for various intensities and colours of the conditioning stimulus.

W. D. Wright[28] has attacked the problem of light-adaptation by using one eye under constant conditions of adaptation, or rather constant conditions of adapting stimulation, as a standard of constant sensitivity with which to compare temporal changes in the sensitivity of the other eye during the course of recovery from exposure to an intense adapting stimulus. By extrapolation to zero time the sensitivity corresponding to the given state of adaptation is deduced. Measurements were made with a special binocular arrangement of Wright's trichromatic colorimeter. Test stimuli of various colours and adapting stimuli of various intensities and colours were used.

Several results of fundamental importance to the theory of the adaptation process are brought out in this paper, one of the most significant being that the response of the photosensitive mechanism of the eye appears to be directly proportional to the stimulus-intensity. A second result of great significance is that when the eye is fully adapted to bright fields the *sensation* tends to reach a constant level for a very large range of stimulus-intensities.

The author also deduces a relation which should enable the fundamental response

functions of the tri-receptor mechanism* to be determined from the results of these adaptation experiments, but there is internal evidence that his first attempt at evaluating these functions is in error.

Y. le Grand[29] has studied the effect on the brightness threshold for a 3° bipartite field of a bright glare source, 5° from the test field, the field brightness ranging from 10^{-6} to 10^{-2} candles per cm^2. The glare source had an angular diameter of 30 min. and a surface brightness of 0·88 cand. per cm^2. Curves are given showing the sensitivity-field brightness relation with and without the glare source and also when a yellow glass is interposed, the main object of the work being to examine the efficacy of yellow filters for reducing the glare of motor-car headlights.

Another important property of vision which has received attention is visual acuity. The most comprehensive study of this matter within the period under review is by R. J. Lithgoe[30]. A statistical method was employed, the test object—Landolt's broken circle—being exhibited to the observer for periods of 1·6 sec. under various conditions as regards size of object, illumination of object, and illumination of surround; and a predetermined proportion of correct recognitions of the position of the break in the circle was taken as the criterion. There are eight possible positions, and the author shows that on theoretical grounds the most suitable criterion to adopt is 4·5 right answers out of eight, and points out that the choice of a different criterion, say seven right out of eight, gives a different relation between acuity and illumination—a consideration generally overlooked.

In general, the visual acuity as defined by the technique employed is found, for the unrestricted pupil, to have a linear relation to the logarithm of the brightness of the test object from the smallest brightness employed (less than 7×10^{-6} cand./cm^2) up to about 4×10^{-3} cand./cm^2, after which the slope of the curve begins to diminish. The effects of surround brightness, the presence of glare sources in the peripheral field and the influence of pupillary diameter are studied. It is shown that with an artificial pupil of fixed size the relation between acuity and the logarithm of the intensity is no longer even approximately linear but is definitely concave to the brightness axis throughout the whole range of intensities investigated.

The paper contains an appendix by E. S. Pearson on the statistical treatment of readings of the kind employed in this work.

M. Kruyswijk and C. Zwikker[31] have studied the relation between visual acuity and brightness, using as a criterion the distance at which the forms of different geometrical solids could be recognized. The tests were made with white light and with sodium light, and the range of brightness was from 5×10^{-5} to 10^{-2} cand./cm^2. The acuity, as defined by the conditions of these experiments, increased steadily

* On the trichromatic theory of vision now widely accepted, the colour and brightness of any visual field are due to the resultant effect of three independent reception systems, whose terminal organs in the retina have different relative sensitivities to radiation of the various wave-lengths comprised within the visible spectrum. Each set of terminal organs is sensitive to a wide range of wave-lengths, but the effect of red and orange light is predominantly due to the response of one set, the effect of green light is predominantly due to the response of a second set and the effect of blue light is predominantly due to the response of the third set. The relative sensitivities of these three sets of receptors, tabulated as functions of wave-length, are what are here meant by "the fundamental response functions of the tri-receptor mechanism."

with brightness over the range studied; the values for sodium light were in general 30 to 40 per cent higher than for white light, but the difference diminished at the higher brightnesses.

The increase in acuity when sodium light is used instead of white light has also been studied by M. Luckiesh and F. K. Moss[32], who used as test objects dark bars on a bright background. The increased acuity obtained with sodium light as compared with light from a tungsten lamp varied from 24 per cent, at $6{\cdot}4 \times 10^{-6}$ cand./cm^2, to 8 per cent at $6{\cdot}4 \times 10^{-3}$ cand./cm^2.

W. De Groot and G. Holst[33] have compared visual acuity in sodium light and in the light from a discharge in zinc vapour which emits chiefly red and blue light. With the latter illuminant acuity was much impaired by the tendency to see coloured fringes bordering the image.

E. Freeman[34] has investigated the effect of the distance of the test object on the power of resolving two vertical black bars on an illuminated background. He finds that acuity increases with distance for a test object subtending a constant angle at the observer, and shows that this is not due to accommodation, as the effect persists when suitable lenses are placed before the eye so as to maintain a constant state of accommodation. No explanation is offered.

M. Luckiesh and F. K. Moss[35] have confirmed Freeman's result.

The behaviour of the eye under intermittent illumination is of great importance in relation to the mechanism of visual reception. The most important contributions to our knowledge of these matters since the work of Ives some 12 to 15 years ago have been made in a series of papers by S. Hecht[36] and his collaborators. Considerations of space prevent any detailed analysis of these papers, which, taken collectively, carry our knowledge of intermittent vision considerably forward. Fresh observations, made with refined measuring apparatus, are presented, and a theoretical discussion of the results is given. On the observational side, the results show a marked increase in precision over previous work and, as would therefore be expected, exhibit features which are absent from previous results. For example, the familiar representation of the relation between critical frequency and the logarithm of the retinal illumination as two straight lines is seen to be only a rough approximation to the actual state of the case, especially at very low intensities. Marked departure from the straight-line law occurs also at high intensities. Results are given for various retinal areas, including different proportions of rods and cones, from which it is possible to identify the features for which these different types of receptor are responsible.

On the theoretical side, it is shown that the fusion of the effects of intermittent stimuli above the critical frequency must take place at the retinal end and not the central end of the neural mechanism; that is to say the intermittency must be absent from the nerve impulse. While admitting that there may be other less obvious factors which modify the behaviour in detail, the authors set themselves the task of finding to what extent the results can be predicted from the kinetics of a reversible photo-chemical reaction. It is shown that, assuming the *first* stages of the visual process to be determined by such a reaction, the validity of Talbot's law, hitherto

regarded as a purely empirical law, can be established, and also that the experimental results of critical frequency measurements can be very closely predicted.

The authors find it unnecessary to take account of the rate of diffusion of the photo-products, as was done by Ives, but obtain formulae which fit the facts very closely from the relation between concentration and reaction velocity alone. Certain minor discrepancies remain to be cleared up, but the evidence of the general adequacy of the theory is not seriously affected by these minor details.

Work on the same problem has also been published by P. W. Cobb[37]. The Ives theory of flicker is criticized on the ground that the effects of change in the stimulus cycle are not correctly reproduced. The validity of Ives' resolution of the square-topped stimulus into its Fourier components, and placing the onus of determining critical frequency relations on the first term of the series, is questioned, as is also the part played in Ives' treatment by the rate of diffusion of the photo-products. The author concludes that the relation between critical frequency and illumination remains an empirical one.

The author also investigates the effect of the ratio of the duration of the light phase to the total period on the critical frequency when the average intensity, as given by Talbot's law, is maintained constant. Certain features of the results are found to be incompatible with the hypothesis that the second, third, etc. terms of the Fourier series contribute jointly to the flicker effect, and the general conclusion is reached that there are difficulties in providing a satisfactory explanation of the physiology of flicker by means of the resolution of the stimulus into sinusoidal components.

Among papers on miscellaneous visual problems, one which is of both theoretical and practical interest has been published by W. S. Stiles and B. H. Crawford[38] on the "Luminous efficiency of rays entering the eye pupil at different points." The results indicate that the apparent brightness of an object is not proportional to the pupil area, as usually assumed, because rays entering the pupil at points off the axis are not so effective in stimulating vision as those which enter axially. Curves are given showing the relative efficiency of different zones of the pupil. Though various hypotheses to account for this are examined, no satisfactory explanation is advanced. The authors point out various circumstances in which this effect may introduce unsuspected errors in certain kinds of measurement. It is suggested, for example, that it invalidates the usual method of calculating retinal illuminations from pupillary areas, in which it has been the universal custom to assume proportionality between these two variables.

J. F. Allen[39] confirms the binocular fusion of complementary colours to make white, and also of red and green to make yellow. The author regards the latter fact as proof that the *sensation* of yellow is compounded of the *sensations* of red and green. This is not, however, proved by the evidence, which merely shows that the fusion does not occur at the retina or at any part of the nerve track peculiar to each eye, but in a central zone common to the neural system of both eyes. The habit of confusing physiological (i.e. physical) activity in the brain centres with psychological activity, as though they were necessarily the same thing, mars much of the reasoning brought

to bear on these sensory problems. The author also discusses the familiar phenomenon of binocular oscillation, in which first the colour seen by one eye and then that seen by the other predominates, and concludes that the oscillations are not due to fluctuations of attention, but to fluctuations in the inhibitory and facilitative neural reflexes.

J. Saidman[40], by the use of suitable filters transmitting only ultra-violet rays of a limited range of wave-length, gives evidence that the retina is sensitive to radiation of wave-length between 3130 and 3022 angstroms. The sensitivity disappears at about 40 years of age.

The complementary relation between pairs of spectral radiations has usually in the past been investigated by *ad hoc* methods. They can, of course, be deduced directly from the locus of the spectrum in the colour triangle, and E. Genberg[41] has computed the wave-length of the complementary pairs throughout the spectrum, and also the proportions in terms of luminosity, and in terms of energy, in which they match an equi-energy "white" for the "standard observer" defined by the Commission Internationale de l'Éclairage*. As this observer embodies the properties of the average eye as determined by Guild and Wright, the calculated results given in this short paper are equivalent in significance to a separate *ad hoc* investigation of the complementary relations for the seventeen observers employed by these investigators.

E. Stoecker[42] has investigated the visibility of red and white test spots, of $1° 48'$ angular diameter, seen against a white background (*a*) with the naked eye, and (*b*) with a red glass in front of the eye, for background brightness ranging from about $2{\cdot}5 \times 10^{-5}$ to $1{\cdot}8 \times 10^{-3}$ cand./cm.^{2}

G. Durup and H. Pieron[43] have investigated the relation between the value of stimulus intensity expressed in ordinary physical scales (but taking as unit for a given quality of stimulus the threshold intensity) and its value when expressed on the so-called "physiological" scales in which the intensity is expressed as the total number of liminal increments encountered in increasing the stimulus from zero to its actual value. The measurements are made for red and blue light (obtained by means of filters) for fields of various sizes and for foveal and parafoveal regions. The results are compared, at various intensity-levels, with the relative intensities required for equality of brightness of the two stimuli with a view to throwing further light on the Purkinje effect and the inverse effect—the Gallisot effect—which is found in the fovea at very low intensities. Several interesting results are obtained, although, as the authors point out, some important points are left unsolved. Among the conclusions reached are the following: (i) The Purkinje effect is entirely explicable in terms of the distinctive functions of the rods and cones, and no appeal need be made to a difference in the rate of rise of sensation with stimulus intensity for the independent reception systems for blue and red. (ii) The Gallisot effect, which is found in the rod-free fovea, must be explained either by a difference in these rates or by an elevation of the foveal threshold for blue by an inhibitory mechanism. (iii) Equality

* Frequently referred to as C.I.E., or—especially in America—denoted by the initials, I.C.I., of the English version of the title.

of brightness for red and blue does not occur for stimuli containing an equal number of just perceptible steps, and doubt is thrown on the legitimacy of Fechner's integration of liminal increments as a basis of sensation scales.

It is also pointed out that the relevant limen in investigations of this type may be not the difference limen for two juxtaposed fields but a limen applicable to a single area of the retina.

M. Polack[44], writing on "Colour vision and its anomalies," concludes that the trichromatic theory does not agree with the facts, and that colour vision is characterized by two factors, the position of maximum luminosity in the spectrum and the extent of the regions of constant hue at the two extremities of the spectrum. He is clearly unfamiliar with modern interpretations of the trichromatic theory, and with modern determinations of the spectral properties of normal vision, which show there is no range of constant hue at the violet end of the spectrum.

J. S. Haldane[45], in a lecture to the Edinburgh Royal Medical Society, has advanced a new theory of vision, in which simultaneous contrast is the primary, if not indeed the only, cause of colour. Strong criticism is offered of the views which the author supposes physicists hold on the perception of light and colour, namely that the intensity and colour of the sensation evoked by a stimulus falling on any region of the retina are determined solely by the physical characteristics of that stimulus, without reference to the condition of stimulation of other parts of the retina. He is at pains to destroy this view by calling attention to various familiar phenomena of colour and brightness contrast, and the effects to be obtained by looking through a blackened tube. In interpreting these phenomena the author overlooks some other important factors which affect the relation between stimulus and sensation, and some of his main conclusions are thereby vitiated.

The scope and volume of the work covered by this brief survey demonstrates that the problems of vision offer a strong attraction to investigators normally associated with various distinct branches of science, who find a common field of interest among these sensory phenomena. If much of their work is not to be rendered ineffective by mutual failure to appreciate what is fundamental in the point of view and in the technique which the several groups bring to these common problems, closer co-operation must be established. At present the average physicist has considerable difficulty in finding easy access to the work of physiologists and psychologists, and even in learning what work has been published, and where. The same presumably applies to workers in the other groups in respect of work carried out by physicists. The present survey is prepared by a physicist: it is doubtless very far from complete, and would have been still less complete had not several authors courteously sent copies of papers which would not otherwise have reached his attention.

§ 4. COLORIMETRY

The application of measurement to colours, or rather to those properties of light-sources and materials on which their colours, as perceived by the normal human eye, depend, definitely entered on a new phase in the autumn of 1931, when the

C.I.E. adopted resolutions[46] which collectively give a rigorous definition of what is to be understood, for technical purposes, by the colour of a light or the colour of a material, and also define a system of units and scales of measurement on which technical specifications of colour, and the results of colour measurement in general, should be expressed.

Prior to this no attempt had been made to secure the general adoption of any of the systems of colour-measurement which have been brought into use. Different systems have been employed not only in different countries but in different industries within the same country. Co-ordination of most of these systems was impossible owing to the lack of a common basis, and even where possible was rarely effected, except by research workers for their own convenience. The technical users of any system in general saw no reason for employing an alternative mode of expression.

Further, as regards the conditions of illumination and the choice of a standard quality of illumination, no common practice existed, though the colour and intensity of the light remitted by any material depends on both of these factors. The first attempt to give unique significance to colorimetric measurements was made in Great Britain by the development at the National Physical Laboratory of a system of colour measurement and specification in which the illumination was fully standardized. This system was in use at the N.P.L. for about 9 years prior to 1931, and had later been adopted by the principal research laboratories of the British colour industries. In the measurement of colour, as in other measurements, purely national standards, though they may serve many useful purposes, are not ultimately a satisfactory solution of the problem. Progress, in science at any rate, depends more and more on international agreement, and efforts were accordingly made in advance of the 1931 meeting of the C.I.E. to reach an agreed basis for an international system. Though attention had been given to the theoretical aspects of colorimetry in many countries, it was principally in Great Britain and the United States of America that the principles had to any extent been crystallized into a definite practical procedure for the standardization of colour measurement. As a result of discussion between representatives of the N.P.L. and the National Bureau of Standards, the basic principles of a system incorporating the main features of the systems used in these two national laboratories were formulated, and were subsequently adopted by the C.I.E.

Thereby for the first time a colorimetric system attained international status by the action of the appropriate international body. A complete description of the system, showing how it was derived from the older system of the N.P.L., has been given in a paper by Smith and Guild[47]. This paper contains an English translation of the official resolutions, all the official tables defining the properties of the standard illuminants and the properties of the standard observer, additional tables of values of the trichromatic coefficients of monochromatic radiation interpolated at wave-length intervals of 1 mμ., and tables to facilitate the computation of colorimetric values from spectrophotometric tests and the properties of each of the standard illuminants. Methods for calibrating trichromatic colorimeters in terms

of the international system also are fully described, and the transformation equations for converting results in the international system to or from the old N.P.L. system are given.

A shorter paper by Judd[48] also contains the official tables, and in addition extensive tables to facilitate the computation of dominant wave-length and colorimetric purity from the standard form of tri-stimulus specification. Transformation equations for conversion to and from the old Optical Society of America system are added.

Recently T. Smith[49] has prepared condensed tables for the computation of colour co-ordinates from spectrophotometric data when, as is common in practice, the spectrophotometric properties of the sample are adequately specified for practical purposes by measurements at intervals of 10 mμ. For the majority of coloured materials these condensed tables give results indistinguishable from those obtained from measurements at the 5 mμ. intervals for which the standard tables are computed.

The adoption of a standard international system naturally caused a halt in the production of papers on the principles of colorimetry, and since then the work of chief importance in this field has been concerned with the improvement of the technique of fundamental standardization, with the development of new instruments, and the applications of colorimetry to industrial problems. In colorimetry, as in other branches of metrology, the maintenance and improvement of fundamental standards involves an enormous amount of steady work of a character which rarely calls for publication: its influence is felt more in the gradual improvement in the standard of accuracy which is normally attainable in measurements of various kinds. Thus although it is known that work is almost continuously in progress on various problems of colorimetric technique in the national laboratories of Great Britain and the United States, and in various laboratories in European countries, very little of this work has yet been described in scientific journals.

R. Davis and K. S. Gibson[50] have recently published particulars of a liquid filter designed for use with a gas-filled lamp at a colour temperature of 2848° K. to produce light which has the same colour, for the C.I.E. standard observer, as the equi-energy spectrum, which is one of the unitary stimuli of the C.I.E. system. While this filter is not one of the fundamental standards of the system, it may nevertheless prove to have important uses in the calibration of colorimetric apparatus. In the *Annual Report of the National Physical Laboratory*, 1933, an improved type of cell for containing liquid filters of the type used in colorimetry is described. The features of this cell embody the experience gained at the Laboratory in preparing and maintaining such filters as close to specification as possible.

The establishment of a standard colorimetric system has naturally stimulated the invention of new instruments and methods of measurement designed to work in relation to it. H. P. J. Verbeek[51] has described a new trichromatic colorimeter in which spectral primaries are used. Mixing is effected by external reflection, internal reflection, and refraction at the faces of a central fixed prism. Detailed information about the performance of the instrument has not yet been published.

R. A. Houstoun[52] has described methods of colorimetry which utilize the spectral locus of the standard observer on the trichromatic colour chart. The principle underlying these methods is the same as that of the "Vector" method described by Guild[53] in 1925, but the measurements can be carried out with an ordinary spectroscope fitted with a suitably cut quartz prism.

Two methods are considered in the paper. In one of these a colour match is used to determine on the standard trichromatic colour chart a straight line passing through the point corresponding to the test colour, and the position of the point on this line is determined from the relative amounts of the spectral components in the matching stimulus, expressed in terms of luminosity. This method contains all the practical and theoretical defects of the monochromatic-plus-white method and is definitely unsuitable for any purpose other than illustrating principles. In the other method considered, two separate colour-matches are made to determine the test colour uniquely by the intersection of two lines on the colour chart. This method has none of the defects referred to, and is very useful for a variety of purposes. The interest of Houstoun's design is that it requires comparatively little modification of an ordinary spectroscope.

A preliminary description of an instrument designed to illustrate the principles of a photoelectric colorimeter for evaluating the components of colour quality directly on the C.I.E. scales is included in a popular exposition of colorimetry given as one of the discourses at the Physical Society's Annual Exhibition in January 1934 by J. Guild[54], and a brief treatment of the principles underlying the design of such apparatus is given as an appendix to the published account of the discourse.

It is of interest to know that arrangements are being made for the comparison of the colour-temperature scales of the National Bureau of Standards and the N.P.L. It is anticipated that other international comparisons of colour-temperature scales will follow.

Another indication of the growing importance of colour-standardization is the publication of a dictionary of colours by the British Colour Council, a body representative of industrial colour-users. This dictionary contains over two hundred ribbons, each giving both a glossy and a matt surface, with the officially accepted name, B.C.C. number and code word for each colour. In another section are notes, historical and otherwise, on the names adopted and on other names that have been used in Great Britain and in other countries, comparisons with other collections, and a specification of the colours on the C.I.E. system with standard illuminant B. These measurements were made by Messrs Adam Hilger with a Guild colorimeter. Particulars of matches made at the Lovibond Laboratory are also given.

It is difficult, in fact impossible, to obtain, through the medium of the scientific journals, information concerning the volume and character of the work which is being carried out in the application of colorimetry to the technical problems of the colour industries. Such work, at any rate in Great Britain, is almost entirely confined to the laboratories of industrial research associations, and the results of the

investigations are not, as a rule, published through the ordinary channels, but are only made available in confidential papers issued to members of the associations. G. F. New[55], late of the Paint Research Station, Teddington, the laboratories of the Research Association of British Paint, Colour and Varnish Manufacturers, gives a brief indication of some of the directions in which colorimetric measurements are being used in connexion with the problems of fading, pigment mixture, etc., but the detailed description and the results of the extensive work carried out in this Laboratory in recent years are not generally available. Though the policy of withholding general publication of results obtained in industrial research laboratories may be justified, there is a danger of attaching the label "confidential" from habit, and much is perhaps so labelled which could be made available to the general scientific public with gain to everyone. Co-operation between scientific workers inside and outside industry is seriously hindered unless the latter can learn what is being done by the former.

§ 5. PHOTOMETRY

Photometric standards and international co-operation in photometry. When one looks back over a short period of time it is very often difficult to see signs of development in any particular branch of science as progress is usually slow. In photometry, however, if one includes the last few years, considerable progress is apparent. Photometry is probably the Cinderella of the exact sciences. Indeed many people consider photometry to be a rather trivial matter somehow or other connected with a grease spot.

The reason why photometry is, or has been, the most inexact of the exact sciences is, of course, that it is based on the determination of the relative brightnesses of two adjacent surfaces by means of an eye which is different from individual to individual. Hence, if the two juxtaposed surfaces differ in colour, the result of the comparison varies from observer to observer. Not only does it vary with the individual but it varies for the same individual if the portion of the retina used for the comparison is changed, and it may also vary with the order of brightness at which the comparison is made. Photometry has also suffered from the defect that as a foundation of its system of units it has had no primary standard of candle-power reproducible from specification.

The unit of candle-power in use in various countries is at present based on groups of carbon filament electric lamps, the candle-power values for which were assigned after international comparisons many years ago. While the changes in these lamps are probably very small, as they are not often used, it is of course possible for them to change. Particularly is there a danger in the standards of candle-power at other colours corresponding to tungsten vacuum and tungsten gas-filled lamps, differing from country to country when there has been no accepted manner of comparing sources having different colours.

The first steps towards placing the science of photometry on a sound basis were taken in 1924 at the meeting of the I.C.I. (formed in 1901 as the International

Photometric Commission for the purpose of establishing methods of photometry of gas mantle lamps). At that meeting the Commission adopted the following resolutions[56]: (i) The I.C.I. recommends for general use as provisional values for the visibility factor the following.... For special cases which concern the ends of the spectrum or particular conditions of field-size intensity these provisional values may be in default. (ii) The I.C.I. recommends the international adoption of a primary standard of light based on the brightness of a black body operated under conditions which are subject to accurate specification. This Commission recommends to the National Laboratories that they take steps—(1) To formulate standard specifications for the construction and the conditions of operation of a black body as a primary standard of light. (2) To establish a definite value for the brightness of a black body used under these conditions, expressed in international candles per square centimetre.

The first resolution adopted what were in effect the values of the visibility factor obtained by Gibson and Tyndall[57] from measurements on 52 individuals by the direct-comparison step-by-step method for a 3° visual field of brightness of 25 millilamberts (equivalent to an illumination of 25 metre-candles on a perfectly diffusing surface of 100 per cent reflection factor) for most of the spectrum. The results of Tyndall and Gibson enabled it to be concluded that a single standard visibility curve can represent the results obtained with the average eye by either the flicker method or the direct-comparison step-by-step method if limited to a small central region of the retina with fairly high brightness of the photometric field. Comparison with the results of other observers led to the conclusion that further investigation could affect the visibility curve but little, except possibly in the end regions of the spectrum.

Although it was probably not realized at the time, the international adoption of a standard visibility curve for the average eye solved the problem of what was the correct procedure for those photometric comparisons in which the result obtained differed as the conditions of the comparison varied. The correct result is logically that which will be obtained by an observer possessing the agreed average eye, and the practical problem confronting the photometrist is to devise methods which, with real observers possessing different visibility curves, will enable the result corresponding to the average or standard observer to be deduced.

The second resolution was adopted because the experiments of Ives[58] on the brightness of a platinum black body at the melting point had demonstrated the reproducibility of such a standard.

Shortly after 1924 intercomparisons of tungsten filament vacuum lamps indicated that discrepancies occurred in the values of candle-power assigned by the different national laboratories to tungsten-filament vacuum lamps. This implied that the unit when transferred to tungsten vacuum lamps was different from the unit maintained in terms of carbon lamps, and it was also evident that for the bluer gas-filled lamps the units were still more in disagreement.

Accordingly in 1927 representatives of the national laboratories of France, Germany, Great Britain and the United States met at Bellagio during a meeting of

the I.C.I., and formulated a plan[59] for determining which was the most satisfactory method of realizing the unit of candle-power for light of tungsten lamp colour from standards embodying the unit in terms of carbon lamps. The plan consisted in each laboratory determining the transmission factor of blue glasses which represented the colour step from carbon filament to tungsten filament lamps by each of three methods, (i) ordinary contrast Lummer-Brodhun photometry, (ii) flicker photometry under the conditions specified by Ives, viz. 2° visual field, 25 millilambert field brightness (25 metre-candles illumination on a diffusing surface of 100 per cent reflection factor), which conditions Ives[60] had shown resulted in substantial agreement with direct comparison photometry under the same field-size conditions, and (iii) by calculation based on spectrophotometric measurements on the glasses, the standard visibility data and the energy-distribution in the spectrum of radiation from a black body of 2080°, the colour temperature of the carbon lamp standards.

The values obtained by the four laboratories are given below[61,62]:

	Lummer-Brodhun	Flicker	Spectrophotometer
N.P.L.	0·5296	0·5205	0·5232
N.B. of S.	0·5238	0·5220	0·5239
P.T.R.	0·5358	0·5280	0·5145
L.C.E.	0·5223	0·509	0·5132

These results were discussed by the representatives of the four laboratories at Cambridge in 1931 during the meeting of the I.C.I. and the following agreement was arrived at: (i) The spectrophotometric results are, in general, the most concordant among themselves. (ii) The spectrophotometric method is certainly correct in principle, provided that the visibility factors have been properly determined. (iii) The agreement between the three methods is sufficiently close, so that no serious practical difficulty will arise from the use of the present visibility factors for measurements on sources of the type now in question. (iv) The laboratories will therefore accept the spectrophotometric method and will proceed jointly to adjust the values of their standard lamps on this basis. (v) This adjustment will make necessary some (small) change in the candle-power of lamps in all countries. This change will be arranged to take place at the same time as the change in electrical units which is expected to be made within a few years, so that all necessary changes in lamp values can be made at one time.

Each of the four national laboratories has been given one of the four blue glasses whose spectral transmission can be considered to be known very accurately from the measurements which have been made on it by each of the four laboratories. It is understood that each of the laboratories is now engaged in setting up new tungsten lamp standards whose candle-power will be determined by means of the blue glass from the existing carbon lamp standards. According to arrangements made, these new standards should be prepared in 1935 and also be intercompared among the various laboratories. It is probable therefore that next year new standards in terms of tungsten vacuum lamps will have been prepared, and if all has proceeded satisfactorily will be adopted and used in the various countries.

It is a property of the Wien expression for black-body or grey-body radiation that a blue glass which represents the colour-change from carbon lamps to tungsten vacuum lamps (a colour temperature change of 2080° K. to 2400° K.) will also represent the colour-change from tungsten vacuum to tungsten gas-filled lamps (a colour temperature change of 2400° K. to about 2850° K.), so that ultimately a somewhat similar procedure can be adopted to bring about uniformity in the unit of candle-power for tungsten gas-filled lamps.

Meanwhile, the International Committee on Weights and Measures has taken up the question of photometry. The International Committee is a body on which 32 nations are represented. It was given power to deal with international electrical units in 1921. In 1927 it was further authorized to set up an Advisory Committee on Electricity which was to include representatives of certain national laboratories. In 1929 the International Committee decided to consider the problem of an international system of photometric units and instructed the Advisory Committee on Electricity to consider the problems of photometry. The Advisory Committee was also authorized to seek the co-operation of national laboratories and also of the I.C.I., which up to 1929 was the only international body interested in such problems.

Photometric problems were discussed at the meeting of the Advisory Committee in 1930 and in 1933. At the 1930 meeting of the Advisory Committee, the Bureau of Standards presented the results of a determination of the brightness of a black body at the freezing point of platinum. Three determinations of this quantity had already been made. That by Ives[58] in which a narrow slit in a hollow cylinder of platinum at its melting point gave a value of 55·4 cand./cm^2 That of Brodhun and Hoffman[63] at the Physikalisch-Technische Reichsanstalt obtained in a furnace in which the appropriate temperature was established by the melting of a piece of platinum wire gave 58·72 cand./cm^2, and that by Wensel, Roeser, Barbrow and Caldwell[64] at the Bureau of Standards gave 58·84 cand./cm^2 The Bureau of Standards also presented a specification of the radiator which was to be used. The specification covered the dimensions of the platinum, the size of the tube to be immersed in the platinum, the purity of the platinum and the method of heating, which was to be by inductive heating from a high frequency electrical circuit. The use of thorium oxide as the refractory for containing the platinum was also recommended.

The Advisory Committee adopted resolutions supporting the opinion of the I.C.I. that a black-body radiator operated under specified conditions should be adopted as the primary standard of light and requested the various national laboratories to examine the specification proposed by the Bureau of Standards. The Advisory Committee also expressed its interest in the work carried out on the transmission values of the blue glasses by the four national laboratories with the object of bringing their candle-power standards into agreement.

At the meeting of the Advisory Committee in 1933 it was decided to recommend that the International Committee on Weights and Measures should separate Electricity and Photometry by setting up a new Advisory Committee on Photometry, the personnel of which should be selected in consultation with the I.C.I., which at its meeting in 1931 had set up a Select Committee on Photometry. The

meeting of the Advisory Committee in 1933 considered reports from the N.P.L.[65] and from Prof. Ribaud of Strassburg[66] on the proposed primary standard of light. Buckley and Barnett at the N.P.L. had obtained a value of 59·1 cand./cm² as the brightness of the black body at the freezing point of platinum. Prof. Ribaud obtained 58·78 cand./cm² and the National Bureau of Standards confirmed their previous value as 58·86.

The Advisory Committee adopted a resolution to the effect that it was premature to adopt a definite value for the standard at the moment. It did, however, agree that the proposed primary standard should be based on a black body at the freezing point of platinum and invited the national laboratories to determine the value of the brightness of the black body at this temperature by 1935.

In addition to being concerned with the primary standard of light, the Advisory Committee discussed heterochromatic photometry and expressed the opinion that the first step in the solution of the problem of heterochromatic photometry was the adoption of a standard visibility curve. The Committee recommended the adoption of the visibility curve already adopted by the I.C.I. and also recommended that the method by which it is to be used should be considered in co-operation with the national laboratories and the I.C.I.

As regards the existing unit of candle-power, the Advisory Committee at its meeting in 1930 recommended that the national laboratories exchange and compare groups of lamps operated at the colour given by carbon filaments. The results of these measurements were communicated to the meeting of the Advisory Committee in 1933. They showed that at this colour the international candle as maintained in France, Great Britain and the United States was the same within the limits of experimental error, while the German Hefner candle appeared to be approximately 0·89 of the international candle as against the nominal value of 0·90 determined in 1909.

This outline of developments mainly in the semi-political field shows that the problems of setting up a reproducible primary standard of light and of standardizing methods by which heterochromatic photometry should be carried out are now being thoroughly investigated by competent international organizations, and it seems likely that in the course of the next few years we shall see the establishment of a scientific system of photometric units and of a standard procedure for heterochromatic photometry.

The general heterochromatic photometry problem. The problem of heterochromatic photometry discussed above has been concerned with the problem mainly as it affects the photometry of sources emitting black- or grey-body radiation. The general problem of heterochromatic photometry is of course far wider and is met with in its most acute form in the photometry of sources emitting radiations approximating to monochromatic radiations. Such sources are met with in the luminous discharge tubes which have been developed in the last few years. The principal gases used in these sources are neon, sodium vapour, and mercury vapour. Paterson[67], Fabry, Roux and Perrin[68], Yamauti[69], and Dziobek and Reeb[70], have described coloured filters

which in conjunction with tungsten lamps colour-match the radiation from the discharge tubes. It is evident, however, that a colour-match does not mean identity of energy-distribution in the two photometric fields which are thus compared. The psychological difficulties in deciding what is a brightness-match are reduced but the inherent differences between observers may still be quite large. The work of Buckley[71] at the N.P.L. has shown[72] how in these circumstances the results are affected by the size of the photometric field (i.e. the portion of the retina used). Where colour filters are used, the light transmissions of the filters are determined by calculation based on spectrophotometric measurements of the filters and the visibility curve of the average eye. Since this latter was established for a field-size of 3° it would be logical to adopt the 3° field as the standard field-size in heterochromatic photometry whenever the experimental conditions are such that a change of field-size makes a significant difference in the results obtained.

The flicker method of photometry was also investigated by Buckley[71] in connexion with the measurements of the integral transmissions of coloured filters and the candle-power of luminous discharge tubes. The results on the luminous discharge tubes showed that values obtained by the Guild flicker photometer were not very different from those obtained by the equality of brightness or Lummer-Brodhun photometer used with coloured filters of calculated transmission. As mentioned above, however, there were certain systematic differences in the results obtained by the equality of brightness and Lummer-Brodhun photometers with small and large fields. As regards the comparison of the transmission ratios of coloured filters by means of calculation and by means of the flicker photometer, the results were in fair agreement with those of Crittenden and Taylor[73] in America who concluded that any difference between the two methods was for practical purposes small when one considered the inherent difficulties in such highly heterochromatic photometry. Winch and Palmer[74] have described methods of determining the suitability of the colour-match filters used in the photometry of discharge lamps.

Physical photometry. Considerable attention continues to be devoted to the use of photoelectric cells for photometry. When the photoelectric cell is used to compare illuminations of the same colour and approximately the same intensity, the fact that the spectral sensitivity curve is different from that of the eye is not of importance. When small colour differences are measured the cells may be made to have sensitivity curves approximating to that of the eye by the use of an appropriate filter (liquid, gelatine or glass). Many laboratories are making use of photoelectric cells in this way and most lamp factories now carry out the rating of their lamps by means of photoelectric photometers in conjunction with integrating spheres. The filter is invariably so designed that the combined photocell and filter gives the same results as are obtained visually. Thus the combination replaces the eye and the values obtained are merely the visual values obtained by the photoelectric method with consequent gain of increased speed of working and high precision in reproducibility of results.

In photoelectric photometry it does not yet seem that the high reproducibility of results attainable by photoelectric methods has enabled these methods to be used

for the highest grade photometry. Many workers have described filters for use with various cells giving approximations to the average eye, but the degree of approximation required in order that the cell thus corrected may give the same results as the average eye even in the comparison of vacuum with gas-filled lamps does not yet appear to have been attained. Still less is it yet possible to use such cells and filters for the comparison of gaseous discharge lamps with tungsten lamps without empirical calibration based on visual measurements.

Winch[75,76] has described methods of using cells which have been very successfully applied in the lamp factory, while Walsh[77] has described methods in use at the N.P.L. based on the work of Harrison, Preston and McDermott. Where valve amplification is used the circuit which appears to be the most successful is that which was used by Sharp and Smith[78]. In this circuit the photocurrent passes through a high resistance, the voltage drop down which is communicated to the grid of a thermionic valve. This valve forms one arm of a Wheatstone bridge of which the other arms consist of two resistances and a second thermionic valve closely matching the first valve in electrical properties.

The bridge is first balanced with the photoelectric cell shielded and then, when the cell is illuminated, again balanced by the application of a neutralizing voltage to the high resistance. The magnitude of the applied voltage is proportional to the illumination if the photoelectric current is proportional to the illumination received by the cell.

Some of the difficulties of photoelectric photometry appear to be inherent in the cells themselves. A comprehensive survey of the behaviour of cells of different types has been given by Ives and Kingsbury[79]. Preston and McDermott[80] have shown that, though the current from cells of the vacuum thin layer type may be strictly proportional to the illumination, in many cells such close proportionality does not hold, and that the departure from proportionality may vary with the wavelength of light used. Thus the photoelectric cell has features analogous to the Purkinje effect in the human eye.

Thin-layer vacuum cells seem to be used for photometric purposes more than any other type. Their sensitivity approaches that of the earlier gas-filled bulk metal cells and they appear to possess greater stability.

Considerable interest has been taken in the last few years in the blocking-layer or *sperrschicht* type of cell, particularly the selenium and the copper-oxide types. A considerable number of papers have been written about the theory and mechanism of such cells but not very much has yet appeared concerning their use in photometry. They are, however, finding a very useful application in portable illumination photometers, the absence of the need for a separate battery being very valuable. One recent model incorporates six of the cells in parallel. It is claimed that the cells are capable of measuring illuminations over the range 0·05 metre-candle to 10,000 metre-candles. Cells of this type exhibit deviations from the cosine law as the angle of incidence of the illumination varies and various methods of correction have been devised.

Recent developments, particularly in Holland, have led to considerably increased

precision in radiometric measurements. In consequence Ornstein[81], at the meeting of the I.C.I. in 1931, suggested that photometric measurements might be based on spectro-radiometric measurements of energy combined with the agreed visibility curve. The possibility of carrying out photometry by the aid of thermopiles corrected by means of a solution which make the spectral response curve of the thermopile approximately identical with that of the average eye was investigated many years ago by Ives. Recent improvements in thermopile and galvanometer construction would appear to warrant further investigation in the subject. In this connexion, Krefft and Pirani[82] have determined the distribution of energy in absolute units of various types of luminous discharge lamps through the whole spectrum from 280 to 15,000 mμ. The measurements were carried out primarily for the purpose of obtaining information about the lamps, but the method if developed could possibly form the basis of a method of physical photometry.

Light-sources. Considerable progress has been made in the last few years in connexion with luminous discharge lamps. These date back to the early years of the century when the Moore tube and Cooper Hewitt lamps were developed. High-voltage neon tubes have been used for several years, particularly in connexion with advertisement signs. The high voltage necessary in the Moore tubes which at first contained air and later nitrogen or carbon dioxide restricted their use, while their efficiency was not very high and was surpassed by that of the gas-filled lamps which were developed about 1912. The application of the hot cathode to the luminous discharge tube by Found and Dorney[83] in America reduced the very high voltage-drop which took place in the neighbourhood of the cathode and recent developments have led to the production of lamps which will operate on voltages of 220 and 110. Of the various sources now used, sodium vapour is the most efficient. A theoretical efficiency of about 360 lumens per watt as against the maximum efficiency of about 620 lumens per watt is possible. The theoretical efficiency has been achieved by Pirani under rather special conditions. In actual practice, however, 70 lumens per watt are obtained. The losses are due to the fact that the sodium vapour becomes hot and heats up the walls of the tube which in consequence radiate in the infra-red. As sodium vapour attacks ordinary glass, the sodium vapour is enclosed in a special resisting glass which is surrounded by a further glass envelope, the space between the two glasses being evacuated to reduce heat-losses. The sodium-vapour lamp is being used abroad in Germany and Holland on motor highways. An experimental system has also been set up in England on part of the Brighton road near Croydon.

The other gases which have been used in commercial lamps are neon and mercury vapour. These gases as well as the sodium vapour are usually mixed with another gas whose effect is to aid the starting of the discharge and only slightly to modify the light emitted, which is mainly that of the neon, mercury or sodium vapour. The luminous efficiency obtained is about 40 lumens per watt for the mercury-vapour lamp in the 400-watt size.

This latter lamp is now becoming fairly extensively used in road-lighting in England where it is stated that, owing to the low level of illumination, colour-

discrimination is very low, so that the colour-distortion produced by the mercury-vapour light is not a serious deterrent to its acceptance by the public. It is claimed that visual acuity is higher with monochromatic light such as that emitted by the sodium-vapour lamp and much interest is being taken in the question.

Attempts are being made to reduce the colour-distortion obtained with mercury-vapour lamps by the addition of zinc and cadmium, which under the conditions of the discharge tube emit lines in the red and blue portions of the spectrum. Such lamps are much kinder to the complexion than the ordinary mercury-vapour lamp but the luminous efficiency is reduced to about 35 lumens per watt.

These mercury-vapour lamps are designed for use on an a.c. supply with a choke, in which case the power factor is about 0·6. This is improved to about 0·9 by the use of a condenser across the mains. An account of the development and present status of luminous discharge tubes has been given by Ryde[84], of the Research Laboratories of the General Electric Co., Ltd.

REFERENCES

(1) A. E. Conrady. *Applied Optics and Optical Design.* Oxford University Press (1929).
(2) *Report on the Teaching of Geometrical Optics.* The Physical Society (1934).
(3) G. Prance. *W. R. Hamiltons Abhandlungen zur Strahlenoptik.* Akademische Verlagsgesellschaft, Leipzig (1933).
(4) H. Boegehold. *Über die Entwicklungen-Theorie der optischen Instrumente seit Abbe.* J. Springer (1929).
(5) M. Berek. *Grundlagen der praktischen Optik: Analyse und Synthese optischer Systeme.* W. de Gruyter & Co. (1930).
(6) M. Herzberger. *Strahlenoptik.* J. Springer (1931).
(7) J. Picht. *Optische Abbildung: Einführung in die Wellen- und Beugungstheorie optischer Systeme.* F. Vieweg & Sohn (1931).
(8) M. von Rohr and H. Boegehold. *Das Brillenglas als optisches Instrument.* J. Springer (1934).
(9) C. Zwikker. *Leerboek der Optik.* P. Noordhoff (1933).
(10) F. Zernike. *Mon. Not. R. astr. Soc.* **94**, 377 (1934).
(11) C. R. Burch. *Mon. Not. R. astr. Soc.* **94**, 384 (1934).
(12) L. C. Martin. *Proc. phys. Soc.* **43**, 186 (1931).
(13) —— *Proc. phys. Soc.* **46**, 231 (1934).
(14) M. Berek. See Münster (15) for numerous references to Berek's papers.
(15) C. Münster. *Ann. Phys.*, Lpz., **15**, 619 (1932).
(16) A. Ames, K. N. Ogle and G. H. Gliddon. *J. opt. Soc. Amer.* **22**, 538, 575, 665 (1932).
(17) A. W. Tronnier and J. Picht. *Z. tech. Phys.* **13**, 402 (1932).
(18) L. W. Tilton. *Bur. Stand. J. Res.*, Wash., **2**, 909 (1929).
(19) —— *Bur. Stand. J. Res.*, Wash., **6**, 59 (1931).
(20) —— *Bur. Stand. J. Res.*, Wash., **11**, 25 (1933).
(21) *Rep. of Phys. and Opt. Societies, Discussion on Vision*, June 1932, Cambridge Univ. Press.
(22) L. C. Martin, F. L. Warburton and W. J. Morgan. *Spec. Rep. Ser. med. Res. Coun.*, Lond., No. 188 (1933).
(23) D. B. Judd. *J. opt. Soc. Amer.* **23**, 1 (1933).
(24) E. P. T. Tyndall. *J. opt. Soc. Amer.* **23**, 15 (1933).
(25) W. D. Wright and F. H. G. Pitt. *Proc. phys. Soc.* **46**, 459 (1934).
(26) S. Hecht, G. Ward and C. Haig. *Arch. Sci. Biol.*, Napoli, **18**, No. 1, 4 (1933).
(27) W. S. Stiles and B. S. Crawford. *Proc. roy. Soc.* B, **113**, 496 (1933).

(28) W. D. Wright. *Proc. roy. Soc.* B, **115**, 49 (1934).
(29) Y. le Grand. *C.R. Acad. Sci.*, Paris, **198**, 1075 (1934).
(30) R. J. Lithgoe. *Rep. med. Res. Coun.*, Lond., No. 173 (1932).
(31) M. Kruyswijk and C. Zwikker. *Physica*, 's Grav., **1**, 225 (1934).
(32) M. Luckiesh and F. K. Moss. *J. Franklin Inst.* **215**, 401 (1933).
(33) W. De Groot and G. Holst. *Physica*, Eindhoven, **13**, 231 (1933).
(34) E. Freeman. *J. opt. Soc. Amer.* **22**, 402, 729 (1932).
(35) M. Luckiesh and F. K. Moss. *J. opt. Soc. Amer.* **23**, 25 (1933).
(36) S. Hecht and C. D. Verrijp. "The influence of intensity, colour and retinal location on the fusion frequency of intermittent illumination," *Proc. nat. Acad. Sci.*, Wash., **19**, 522 (1933).
S. Hecht and E. Wolf. "Intermittent stimulation by light. I. The validity of Talbot's law for Mya," *J. gen. Physiol.* **15**, 369 (1932).
S. Hecht, S. Shlaer and C. D. Verrijp. "Intermittent stimulation by light. II. The measurement of critical fusion frequency for the human eye."
S. Hecht and C. D. Verrijp. "Intermittent stimulation by light. III. The relation between intensity and critical fusion frequency for different retinal locations."
—— —— "Intermittent stimulation by light. IV. A theoretical interpretation of the quantitative data of flicker."
Papers II, III and IV are printed together in *J. gen. Physiol.* **17**, 237–82 (1933).
(37) P. W. Cobb. *J. opt. Soc. Amer.* **24**, 91, 107 (1934).
(38) W. S. Stiles and B. H. Crawford. *Proc. roy. Soc.* B, **112**, 428 (1933).
(39) J. F. Allen. *Canad. J. Res.* **8**, 390 (1933).
(40) J. Saidman. *C.R. Acad. Sci.*, Paris, **196**, 1537 (1933).
(41) E. Genberg. *Proc. phys. Soc.* **45**, 836 (1933).
(42) E. Stoecker. *Beitr. Geophys.* **40**, 75 (1933).
(43) G. Durup and H. Pieron. *Année psychol.* 57 (1932).
(44) M. Polack. *C.R. Acad. Sci.*, Paris, **197**, 1003, 1253 (1933).
(45) J. S. Haldane. *Brit. med. J.* Dec. 23rd, 1933, p. 1153.
(46) *Comm. Int. de l'Eclair., Compte Rendu*, VI^e Sess., p. 19 (1931).
(47) T. Smith and J. Guild. *Trans. opt. Soc.* **33**, 98 (1931–2).
(48) D. B. Judd. *J. opt. Soc. Amer.* **23**, 359 (1933).
(49) T. Smith. *Proc. phys. Soc.* **46**, 372 (1934).
(50) R. Davis and K. S. Gibson. *Bur. Stand. J. Res.*, Wash., **12**, 263 (1934).
(51) H. P. J. Verbeek. *Physica*, Eindhoven, **13**, 77 (1933).
(52) R. A. Houstoun. *Phil. Mag.* **16**, 945 (1933).
(53) J. Guild. *Trans. opt. Soc.* **27**, 139 (1925–6).
(54) J. Guild. *J. sci. Instrum.* **11**, 69 (1934).
(55) G. F. New. *Trans. ceram. Soc.* **32**, 95 (1933).
(56) *Comm. Int. de l'Eclair., Compte Rendu*, VI^e Sess., p. 67 (1924).
(57) K. S. Gibson and E. P. T. Tyndall. *Bull. U.S. Bur. Stand.* **19**, 131 (1923).
(58) H. E. Ives. *J. Franklin Inst.* **197**, 147, 359 (1924).
(59) *Rep. nat. phys. Lab.* p. 130 (1927).
(60) H. E. Ives. *Phil. Mag.* **24**, 853 (1912).
(61) *Comm. Int. de l'Eclair., Compte Rendu*, VIII^e Sess. (Cambridge), p. 114 (1931).
(62) *Comm. Int. des Poids et Mes.*, Pro. Verb., **16**, 296, 323 (1933).
(63) E. Brodhun and F. Hoffman. *Z. Phys.* **37**, 137 (1926).
(64) H. T. Wensel, W. F. Roeser, L. E. Barbrow and F. R. Caldwell. *Bur. Stand. J. Res.* **6**, 1103 (1931).
(65) *Comm. Int. des Poids et Mes.*, Pro. Verb., **16**, 256 (1933).
(66) *Comm. Int. des Poids et Mes.*, Pro. Verb., **16**, 261 (1933).
(67) C. C. Paterson. *Illum. Engr*, Lond., **25**, 308 (1932).
(68) C. Fabry, L. Roux and E. Perrin. *Rev. Opt.* (*théor. instrum.*), **8**, 1 (1929).
(69) Yamauti. *Comm. Int. de l'Eclair., Compte Rendu*, VIII^e Sess. (Cambridge), p. 188 (1931).
(70) W. Dziobek and O. Reeb. *Z. Phys.* **14**, 350 (1933).
(71) Buckley. *Illum. Engr*, Lond., **22**, 118, 148 (1934).

(72) Since this report was written a paper on this subject has been published by WINCH, PALMER, MACHIN and DUDDING. *G.E.C. Journ.* p. 125 (1934).
(73) E. C. CRITTENDEN and A. H. TAYLOR. *Trans. Illum. Engng Soc.*, N.Y., **24**, 153 (1929).
(74) G. T. WINCH and E. H. PALMER. *Illum. Engr*, Lond., **27**, 223 (1934).
(75) G. T. WINCH. *J. sci. Instrum.* **6**, 374 (1929).
(76) —— *J. Instn elect. Engrs*, **68**, 533 (1930).
(77) J. W. T. WALSH. *Illum. Engr*, Lond., **26**, 64 (1933).
(78) C. H. SHARP and H. A. SMITH. *Trans. Illum. Engng Soc.*, N.Y., **23**, 428 (1928).
(79) H. E. IVES and E. F. KINGSBURY. *J. opt. Soc. Amer.* **21**, 541 (1931).
(80) PRESTON and MCDERMOTT. *Proc. phys. Soc.* **46**, 256 (1934).
(81) L. S. ORNSTEIN. *Comm. Int. de l'Eclair., Compte Rendu*, VIII^e Sess. (Cambridge), p. 139 (1931).
(82) H. KREFFT and M. PIRANI. *Z. tech. Phys.* **14**, 393 (1933).
(83) C. G. FOUND and J. G. DORNEY. *J. Amer. Inst. elect. Engrs*, **47**, 855 (1928).
(84) N. RYDE. *J.R. Soc. Arts*, **82**, 624 (1934).

HEAT

By J. H. AWBERY, B.A., B.Sc., F.Inst.P., National Physical Laboratory

§ 1. INTRODUCTORY

The chief changes which have taken place in the subject of heat in recent years are traceable on the theoretical side to the quantum theory, and on the experimental side, to the application of electricity.

When the writer was at school, not such a very long time ago, he remembers well the distinction that was made between the determination of the mechanical equivalent of heat and that of the electrical equivalent. No reason now exists for maintaining this distinction, and heat quantities are perhaps more frequently measured electrically than by the aid of a water calorimeter.

Of course, measurements imply units, and these have within quite recent years, changed their nature appreciably.

Not only has the unit of heat been overhauled by International Committees, but a scale of temperature has been accepted by the majority of standardizing laboratories. Until 1931 this scale defined temperature between 0° and 100° C. by means of the mercury thermometer, but this has now been dropped, and extracts from the present form of the scale (which agrees as nearly as is known with the thermodynamic scale) are shown below.

"The basic fixed points and the numerical values assigned to them for the pressure of one standard atmosphere are given in table 1, together with formulae which represent the temperature (t_p) as a function of vapour pressure (p) over the range 680–780 mm. of mercury.

Table 1. Basic fixed points of the international temperature scale.

(*a*) Temperature of equilibrium between liquid and gaseous oxygen at the pressure of one standard atmosphere (oxygen point) ... ... −182·97° C.

$$t_p = t_{760} + 0{\cdot}0126\,(p-760) - 0{\cdot}0000065\,(p-760)^2$$

(*b*) Temperature of equilibrium between ice and air-saturated water at normal atmospheric pressure (ice point) ... ... 0·000° C.

(*c*) Temperature of equilibrium between liquid water and its vapour at the pressure of one standard atmosphere (steam point) ... ... 100·000° C.

$$t_p = t_{760} + 0{\cdot}0367\,(p-760) - 0{\cdot}000023\,(p-760)^2$$

(*d*) Temperature of equilibrium between liquid sulphur and its vapour at the pressure of one standard atmosphere (sulphur point) ... ... 444·60° C.

$$t_p = t_{760} + 0{\cdot}0909\,(p-760) - 0{\cdot}000048\,(p-760)^2$$

(*e*) Temperature of equilibrium between solid silver and liquid silver at normal atmospheric pressure (silver point) ... ... 960·5° C.

(*f*) Temperature of equilibrium between solid gold and liquid gold at normal atmospheric pressure (gold point) ... ... 1063° C.

Standard atmospheric pressure is defined as the pressure due to a column of mercury 760 mm. high, having a mass of 13·6961 gm./cm^3, subject to a gravitational acceleration of 980·665 cm./sec^2 and is equal to 1,013,250 dynes/cm^2.

It is an essential feature of a practical scale of temperature that definite numerical values shall be assigned to such fixed points as are chosen. It should be noted, however, that the last decimal place given for each of the values in the table is significant only as regards the degree of reproducibility of that fixed point on the International Temperature scale. It is not to be understood that the values are necessarily known on the Thermodynamic Centigrade scale to the corresponding degree of accuracy.

The means available for interpolation lead to a division of the scale into four parts.

From the ice point to 660° C. the temperature t is deduced from the resistance R_t of a standard platinum resistance thermometer by means of the formula

$$R_t = R_0 (1 + At + Bt^2).$$

The constants R_0, A and B of this formula are to be determined by calibration at the ice, steam and sulphur points respectively.

The purity and physical condition of the platinum of which the thermometer is made should be such that the ratio R_t/R_0 shall not be less than 1·390 for $t = 100°$ and 2·645 for $t = 444·6°$.

From $-190°$ *to the ice point* the temperature t is deduced from the resistance R_t of a standard platinum resistance thermometer by means of the formula

$$R_t = R_0 (1 + At + Bt^2 + C (t - 100) t^3).$$

The constants R_0, A and B are to be determined as specified above, and the additional constant C is determined by calibration at the oxygen point.

The standard thermometer for use below 0° C. must in addition have a ratio R_t/R_0 less than 0·250 for $t = -182·97°$.

From 660° C. *to the gold point* the temperature t is deduced from the electromotive force e of a standard platinum *vs* platinum-rhodium thermocouple, one junction of which is kept at a constant temperature of 0° C. while the other is at the temperature t defined by the formula

$$e = a + bt + ct^2.$$

The constants a, b and c are to be determined by calibration at the freezing point of antimony, and at the silver and gold points. The freezing point of the antimony employed is to be determined by use of the resistance thermometer calibrated as prescribed above.

Above the gold point the temperature t is determined by means of the ratio of the intensity J_2 of monochromatic visible radiation of wave-length λ cm., emitted by a black body at the temperature t_2, to the intensity J_1 of radiation of the same wave-length emitted by a black body at the gold point, by means of the formula

$$\log_e \frac{J_2}{J_1} = \frac{c_2}{\lambda}\left[\frac{1}{1336} - \frac{1}{(t+273)}\right].$$

The constant c_2 is taken as 1·432 cm. degrees. The equation is valid if $\lambda(t+273)$ is less than 0·3 cm. degrees."

The full specification of the International Temperature scale contains valuable hints on experimental procedure, and an appendix of secondary fixed points. The latter is transcribed in table 2.

Table 2. Secondary points.

Temperature of equilibrium between solid and gaseous carbon dioxide	−78·5° C.
$t_p = t_{760} + 0{\cdot}01595\,(p-760) - 0{\cdot}0000111\,(p-760)^2$	
Temperature of freezing mercury	−38·87° C.
Temperature of transition of sodium sulphate	32·38° C.
Temperature of condensing naphthalene vapour	217·9$_6$° C.
$t_p = t_{760} + 0{\cdot}058\,(p-760)$	
Temperature of freezing tin	231·8$_8$° C.
Temperature of condensing benzophenone vapour	305·9° C.
$t_p = t_{760} + 0{\cdot}063\,(p-760)$	
Temperature of freezing cadmium	320·9° C.
Temperature of freezing lead	327·3° C.
Temperature of freezing zinc	419·4$_8$° C.
Temperature of freezing antimony	630·5° C.
Temperature of freezing copper in a reducing atmosphere ...	1083° C.
Temperature of freezing palladium	1555° C.
Temperature of melting tungsten	3400° C.

It will be noticed that the scale is not defined below −183° C. This was unimportant whilst Leiden had the only cryogenic laboratory of real importance, but now that the German Reichsanstalt is well established in this field, and both Oxford and Cambridge have set up helium liquefaction plants, the extension of the scale will very shortly become a matter of considerable urgency. This is especially seen when we reflect that a new era is dawning in the low temperature field. When temperatures are lowered by reducing the pressure of liquid helium, a limit is soon reached because a large factor in the pressure makes but a small difference to the temperature. When the demagnetization technique was tried, the first attempt brought the temperature to a point only one-third as far from absolute zero as the previous record. A few weeks later, a temperature of 0·085° K. was reported. The measurement of these temperatures will be discussed later.

Meanwhile, we return to the subject of units. Once the meaning of a degree has been exactly defined, the heat quantity represented by a calorie at any specific temperature becomes definite. Curiously enough, however, just when this stage was reached, electrical methods were so widely used and of such accuracy, that there was a vigorous movement to abandon the calorie altogether. Whether it will ever disappear is for the future to say, but at present it leads a shadowy existence; although still intended to represent the amount of heat required to raise 1 gm. of water 1° C., yet its equivalent in terms of electrical energy is laid down as part of the definition by one body. Should the value of Joule's equivalent be altered, it is not clear what would be done. To use an analogy from statics, we have a self-stressed structure, and some elastic deformation may be necessary in the future.

Another calorie, adopted by the International Union of Pure and Applied Chemistry, is used only for the purpose of standardizing bomb calorimeters used in determining heats of combustion. For this purpose, they have laid down that the heat of combustion of benzoic acid is 6324 gm.-cal./gm. or 26,466 joules/gm. The first-mentioned figure is clearly a definition of the calorie (the 1922 International Combustion Calorie) and need not be changed if new determinations of the heat of combustion were to give a different value in terms of the usual calorie. But it is not clear whether the second figure is a definition of the "1922 International Combustion Joule," or whether it serves to show the relation between the two calories.

§ 2. MEASUREMENT OF TEMPERATURE

The mercury thermometer still remains the most convenient instrument for many purposes. The addition of an inert gas above the mercury column has enabled its range to be increased perceptibly and made the instrument more convenient in that the column is less liable to break. Thermometers with silica envelopes and bulbs have been tried for high temperature work, and are now commercially obtainable, whilst new glasses have been developed which are suitable for the bulbs and stems of mercury thermometers to be used at high temperatures.

The mercury in steel thermometer, with a pointer actuated by a Bourdon gauge and moving over a dial, is now familiar, and has simplified the problem of temperature measurement in industry, where a large instrument with bold figures is often desirable.

Recent years have also seen considerable developments in the use of thermocouples, on which the various catalogues of the Annual Exhibition of the Optical and Physical Societies are a mine of information.

Turning to low temperatures, a special technique has had to be developed. There, as elsewhere, the fundamental standard of reference is the thermodynamic scale, and is most accurately realized by using a gas thermometer. The latter gives directly a good approximation to the true scale, and in addition permits its own deviations from that scale to be calculated if the properties of the gas are themselves known.

Even below the liquefying point of helium, it is possible to use the gas thermometer, provided the pressure of the gas in it is kept so low that it is, throughout the range concerned, below its vapour pressure. The vapour pressure of helium at the lowest temperatures reached before about 1932 is of the order of 0·015 mm. of mercury, and by the use of a constant volume helium thermometer in which the pressure was below this at the corresponding temperature, Kammerlingh Onnes and other workers at Leiden have determined the vapour pressure curve of helium itself. Once this has been carried out with sufficient accuracy, vapour pressure measurements can be used as a convenient thermometer, or resistance thermometers can be calibrated and used for temperature measurement. At very low temperatures (below $-250°$ C.) constantan is found to be a convenient material.

As we rise in the scale, lead becomes a suitable element for the resistance thermometer, and by the time that the boiling point of hydrogen is reached, we find agreement of a most gratifying kind between the measurements of different laboratories. It should be mentioned that below $-183°$ C. the Leiden workers do not entirely agree with some of the other cryogenic laboratories, but use a scale of their own, generally known as the Leiden scale.

The boiling point of liquid oxygen is the lowest point in the International Temperature scale, and above it difficulties are greatly minimized.

For temperatures below the lowest at which the gas thermometer has been used (say $1{\cdot}2°$ K.) recourse must be had to extrapolation of the vapour pressure curve or of the resistance temperature relation. Aid in the former extrapolation is to be had from the law of corresponding states and from the use of well-known forms of vapour pressure law. In the latter case, theory has quite recently had much to say on the question of the way in which electrical resistance should vary with temperature near absolute zero.

These methods are probably fairly reliable, but like all extrapolations, unforeseen circumstances can upset them. Thus, it was long after the introduction of these methods that Keesom [32, 33] discovered the fact that liquid helium exists in two forms, between which there is a transition point, varying with pressure. Fortunately the break in the slope of the vapour pressure curve is only slight, but it illustrates the dangers of the method. The latest advances into the low temperature region have been made by the demagnetization method, and have brought with them their own difficulties in connexion with temperature measurement. The estimates so far made have been based on measurements of the mechanical force experienced by the body in the magnetic field.

To turn to an entirely different aspect of the subject, it may be noted that the exact measurement of flame temperatures has come into some prominence recently. A method analogous to optical pyrometry has been worked out, in which the reversal of spectral lines by means of a background of variable brightness is used. When the background matches the line from the flame, its black body temperature is equal to that of the flame. This method has been extensively used at the American Bureau of Mines, and an account of work with it and another method will be found in the paper by Griffiths and Awbery [20].

The measurement of gas temperatures, even when the gas is not incandescent, is of great importance and no little difficulty, but the conclusions of recent workers are not easy to summarize. The main point is to avoid radiation exchange with neighbouring objects, either hotter or colder than the gas.

For solid objects at higher temperatures, pyrometers of various kinds are in common use, and it is now possible with ordinary commercial instruments to measure a temperature of 1800° C. with an accuracy of 10° C. or so. At the standardizing laboratories, agreement between those of different countries is of the order of 2° at 2000° C. (At temperatures up to 1000° C. a recent intercomparison with thermocouples [5] showed agreement to $0{\cdot}1°$ C. between different laboratories.)

High temperatures are most accurately measured with optical pyrometers based

essentially on Wien's law of radiation and employing monochromatic radiation. The total radiation type of pyrometer based on Stefan's law is also much used, the main advantage of this type being that it allows continuous records of temperature to be made, a feature of great importance in some industrial operations. The accuracy attainable in these high temperature regions is also well illustrated by the fact that determinations of the melting point of platinum at the National Physical Laboratory[4] and the Bureau of Standards[2] disagreed by only 0·2° C. at a temperature of 1773° C.

On the experimental side, one of the difficulties of high temperature work is still the provision of suitable refractory materials. For temperatures up to 1600° C. silica has always been recognized as extremely valuable, but the art of "glass blowing" in silica has much improved in the last 10 or 15 years, so that objects of complicated shape can now be made to quite close limits of measurement. The technique of manufacturing articles in highly refractory materials such as alumina, magnesia, zirconia and thoria, has also been much improved. The use of such articles together with the high frequency induction furnace enables temperatures as high as 2400° C. to be readily attained.

§ 3. CALORIMETRY

The methods of calorimetry have altered but little in the last thirty years. Results are still obtained by the method of mixtures, the method of cooling, electrical methods or by the use of changes of state.

Taking these in order, the method of mixtures has been studied in detail by W. P. White, whose book[19] on the calorimeter is a classic. The old bugbear of the method, the correction for radiation, is now very frequently eliminated by the use of a jacket maintained at the temperature of the calorimeter throughout the whole of the experiment. The necessary temperature change in the jacket may be obtained by electrical heating, by mixing two solutions with a large heat of reaction, or by keeping a flow of water through the jacket, the supply for this flow being from two tanks, one hot and one cold; when this method is used, a valve controls the proportions in which water from the two sources is mixed for delivery to the jacket.

The greatest sources of error to be guarded against nowadays are loss of water by evaporation, and conduction out of the calorimeter along such metallic paths as the shaft of the stirrer, or thermocouple leads. One enthusiast for the adiabatic method has gone so far as to heat his jacket electrically, using automatic apparatus to control the current supply to the jacket, through the medium of differential thermocouples with one junction in the calorimeter and the other in the jacket. These thermocouples conduct no heat out of the calorimeter, since their two ends are at the same temperature. No other temperature-measuring device enters the calorimeter, the readings being taken in the jacket instead, since this is presumably at a temperature equal to that to be measured.

A variant of the method of mixtures, which some would classify as an entirely separate type, is the continuous flow method[8,9], which was developed to an

extremely high degree by Callendar. In this method, it is essential that the heat to be measured should be developed at a steady rate, so that if, for example, the specific heat of a liquid is required, provision must be made for a stream of it to pass through a heater before it enters the calorimeter. This heat is absorbed in a stream of water, of which the permanent temperature rise is measured, either by thermometers, differential thermocouples or resistance thermometers as may be desired. The method has the advantage that temperatures are not changing, so that the heat capacity of the apparatus does not enter the equations. Heat loss from the water stream, during its passage from inlet to outlet, is usually small, but must be taken into account in experiments with any pretension to accuracy.

The best known calorimetric method using a change of state phenomenon for heat measurement is the Bunsen ice calorimeter. Some fifteen years ago, this seemed likely to come into great favour, several authors expending considerable pains to improve its details, and to make it convenient and accurate. It appears now to have subsided again almost to the position of a scientific curiosity.

The same principle has, however, been used in other circumstances. Thus Dewar's pioneer work on the specific heats of metals at low temperatures was carried out by measuring the quantity of liquid oxygen evaporated by the specimen in cooling through a known temperature range. In all experimental work based on such methods, it is necessary to know the latent heat of the substance melted or evaporated; electrical methods give a very convenient means of determining this. It is possible that the method will come more into favour in future, since, among other advantages, it renders a knowledge of the heat capacity of the calorimeter superfluous, owing to the fact that its temperature does not alter.

The attention of the Physical Society has been called to the advantages of cooling methods, in two distinct forms, in papers presented during the past year. In one case[14], the law of cooling is taken to be the $\frac{5}{4}$-power law, and in the other[15], matters are so arranged that Newton's law holds to a high degree of accuracy. Each is convenient in its own range, the former for work of high accuracy on fluids, the latter for a somewhat lower accuracy, but applicable also to solids.

In dealing with calorimetry, the great strides made during this century in gas calorimetry must not be overlooked. Towards the end of the nineteenth century, Thomsen and Berthelot were working at this subject, expending a great deal of labour to obtain results which later work has shown to be of an accuracy of only 2 or 3 per cent at the best. To-day, in every gasworks in the country, heats of combustion are regularly measured with a precision of $\frac{1}{2}$ per cent, in calorimeters which are capable of repeating their results to an even higher precision. A great part of this advance is due to that prince of experimentalists, C. V. Boys, and the complete story may be read in the book by Hyde and Mills[13].

Calorimeters may of course be devised for various special purposes, and one designed by Algera[10] for the measurement of the heat production of plants may be of interest. This heat is of course produced very slowly, and heat losses from the calorimeter during the time required for the attainment of a reasonable temperature rise would be relatively enormous. Consequently, the calorimeter is always main-

tained at a constant temperature, by the ingenious device of pumping air through an immersed tube of distilled water. This cools the calorimeter (which is well stirred) to a point just sufficient to keep its temperature constant, the switching on and off of the pump being controlled automatically by thermocouples in the calorimeter. Since the cooling produced is at any stage just sufficient to compensate the heating due to the plant metabolism, it follows that the total amount of cooling (which is measured by the total volume swept through by the pump piston) measures the total heat generation. Calibration of the pump in terms of heat generated is effected by using electrical heaters in place of the plant material.

§ 4. THE MEASUREMENT OF THERMAL CONDUCTIVITY

In a book—a good one—on Heat, published only eighteen years ago, the methods described for determining thermal conductivity were (1) Forbes' method, and (2) comparative methods.

Twenty-five years ago, an author could assert that he did not believe the conductivity of any gas had really been measured, although he knew publications claiming to describe such measurements.

These instances illustrate the fact that thermal conductivity determinations have progressed more rapidly than any other among the standard measurements in the subject of heat.

In a large measure, the problem is that of calorimetry, and electricity has contributed much to its solution. When we remember the definition of conductivity —heat flowing per unit area in unit time, under unit temperature gradient—it is clear that the two essential parts are the measurement of a heat quantity, and that of a temperature gradient.

When the specimen is in the form of a long rod, there is necessarily a lateral escape of heat, and Forbes was almost compelled to use this escape as his measuring instrument, by investigating the rate of cooling of the rod.

Electrical heating now permits us, if we wish, to jacket the rod with a metal tube, provided with a number of separate heaters adjusted so as to maintain equality of temperature between the tube and the sample throughout their whole length.

In any case, this heat loss is now far less troublesome than it used to be; with thermocouples, which occupy so much less space than thermometers, a gradient can be measured over a short length near the hot or the cold end of the rod, and the energy flowing at these two ends can be measured, in the first case by knowledge of the electrical energy dissipated, and in the second case by collecting it in water flowing in a coil of metal tubing. In Forbes' day the second only of these measurements would have been possible.

When a rod is not a suitable form of specimen, the slab and the disc are available. In these cases, the lateral loss becomes an edge loss, and is usually extremely small in comparison with the heat flowing directly from one face to the other of the specimen. It can be practically eliminated by use of the so-called guard ring method.

In this, one face is heated by an electrical hot plate, which is made in two sections, a central portion and an outer one, with separate windings. The test specimen is just so large as to extend to the outer edge of the outer section, and although it is in one piece, yet the working area is considered to be that part covered by the centre section.

The currents in the two windings are so adjusted that both sections of the hot plate are at the same temperature, when it is clear that the heat flow in the central section (the working area) of the specimen will be normal to the plate, with no side flow. The edge loss is in this method transferred to the outer section, on which no measurements are made.

If the edge loss is not eliminated in this way, it can always be determined experimentally by using as test specimen, a material of which the conductivity is accurately known. The heat which would pass through it if the lines of flow were normal to the plate can then be calculated, the difference between this and the observed heat flow representing the lateral loss.

Two other means of utilizing a disc differ only in the method used for the calorimetry[1, 9]. Both have been described by Prof. Lees. In one, a long metal bar is divided by a cut perpendicular to the axis, and the specimen inserted between the two portions. The temperature in the metal bar is measured by the aid of thermo-couples, at a number of points along its length, in both sections. By extrapolation, the temperatures at the two surfaces of the disc are known, whilst the slopes of the curves show the gradients of temperature on each side of the disc. The product of the gradient and the known conductivity of the metal gives the heat flow per unit area into and out of the disc.

The other method referred to is quite similar, but the heat is allowed to escape from the exposed end of the bar, and its amount is deduced from the temperature difference between the exposed end and the air. This apparatus requires calibration to find the relation between temperature excess and heat loss, but is then very rapid in action.

Another form of specimen which lends itself to accurate determination is the hollow cylinder. The theory is simple, the heat can be generated electrically in an axial heater, and end corrections can be avoided by extensions of the tube at each end, preventing longitudinal flow and acting in a similar manner to the guard ring of a slab apparatus.

The cylinder and the heater can also be amalgamated into one body, in the case of metal specimens, when we return to the problem of the solid rod in a new guise. Instead of heating one end and cooling the other, the heat is now generated throughout the whole volume of the specimen. This method has advantages experimentally, and was used by Callendar[2], and also in the famous investigations of Jaeger and Diesselhorst[3]. A variant of it, the radial gradient being the quantity measured, was described by Angell.

The theory of thermal conduction is not in a very sound position. Consideration will be resumed at a later stage, but it may be noted here that some guidance is obtained in the case of good conductors by considering, not the conductivity itself,

but the Lorenz function, thermal conductivity divided by the product of electrical conductivity and absolute temperature.

The methods briefly mentioned above refer specifically to the problem of determining the thermal conductivity of a solid, which can be brought into the form of a rod, disc or tube. For fluids, whether liquids or gases, similar methods are available, but the precautions necessary to obtain accuracy are greatly increased, since the actual conductivity is usually extremely small. In addition, the thickness of the sample of fluid must be kept so small that convection is suppressed. Whether this result has been achieved can only be tested, in the case of a liquid, by varying the thickness and temperature gradient, and ensuring that this does not affect the result. In a gas, a delicate proof of the absence of convection is provided by the apparent conductivity being independent of the pressure.

§ 5. RADIATION

The theory of radiation was established in its main essentials in the great days of classical physics, when Kirchhoff and Boltzmann were active contributors. The details of the theory had to wait for—indeed were the very origin of—the quantum theory.

As regards black-body radiation, we now know that Stefan's fourth power law is of as high an accuracy as any in physics; we also believe that we know the distribution of energy among the wave-lengths in the spectrum, but this belief is not based on such strong grounds as our belief in Stefan's law. Nevertheless, none of the modern theories has yet suggested that Planck's equation requires modification, though Nernst on experimental grounds once put forward the view that it is only an approximation to the true law. This suggestion was the inspiration for a very exact series of measurements by Rubens and Michel[3], which confirmed the Planck law, and thus presumably disposed of the doubt. G. N. Lewis has also shown[5] how a slightly generalized form of Planck's law can be obtained, if we renounce the assumption that every state of thermal equilibrium is determined by a single variable such as temperature or energy density.

So far, we have been considering mainly black-body, or full, radiation, such as is set up inside any closed system at uniform temperature. The necessity for the latter condition is sometimes overlooked and leads to very confused ideas. Thus, in the space between two concentric brass tubes at the same temperature, the radiation is black-body radiation, and is independent of the emissivity of the walls. But this is not true if the two tubes are at different temperatures. It is not then true that one emits heat at the rate σT_1^4 per unit area and the other returns it at the rate σT_2^4, where σ is Stefan's constant. If the internal one has emissivity E_1 and the other emissivity E_2, each referred to a black body as unity, then in fact the transfer is $\left(\frac{E_1 E_2}{E_1 + E_2 - E_1 E_2}\right)(T_1^4 - T_2^4)\,\sigma$ per unit area of the inner tube, when account is taken of multiple reflections. It may be noticed that if either tube is black, then the first factor reduces simply to the emissivity of the other tube. This is in harmony

σ, E_1 E_2

with the fact that if the outer tube is made infinitely large in diameter (so that it is in fact non-existent and cannot reflect), it must then behave as if it were black. And if it is non-existent, then, by the definition of emissivity, the net heat emission from the inner tube is $E_1 (T_1^4 - T_2^4) \sigma$.

The geometrical calculations in problems of this type are often very involved, and it may be of use to note that results calculated for quite other purposes than this are available in such text-books as Hermann's *Geometrical Optics* and can frequently be transferred bodily to the present subject.

When we consider real bodies instead of the ideal black body, little is known. The total radiation does not accurately follow a power law, though for platinum a good approximation is to take it proportional to the fifth power. The distribution of energy is naturally still more complicated, since every body is "coloured," though its main absorption bands may lie outside the visible region. Even in regions remote from absorption bands, it is not found possible to express the radiation from metals by any formula yet devised.

Not much experimental work has been done recently in this subject, and it is a field which would perhaps repay cultivation.

The laws of the black body have all been applied in temperature measurement, under which heading reference was made to them, but it may be noted here that, if applied to a body which is not black, they all determine an "equivalent" temperature, which is not the true temperature but the temperature of the black body which would have (*a*) the same total emission as the actual body, or (*b*) the same wave-length for maximum emission, or (*c*) the same intensity of emission at the wave-length used, according as the pyrometer is of (*a*) the Féry type, (*b*) the Wien type, or (*c*) the Holborn type.

§ 6. CONVECTION

We have seen that in the last thirty years the theory of conduction has suffered but little change, the advance being on the experimental side, and the theory of radiation has been set on a firm basis, the experimental work being almost confined to measurements of the fundamental constants for a black body. In the case of convection, it would be almost true to say that no experimental work had been done until the last two decades, and that there was no theory until the same epoch. Even now, however, the true theory waits on development in hydrodynamics. For the present we have to be satisfied with the partial theory which arises by the application of the principle of similitude, together with such hints as may be derived from the mathematical investigations of heat transfer into a fluid with stream line motion, under various boundary conditions, some of which are rather artificial. An important paper from this point of view is the one by Piercy and Winny[19].

Another method which has given quite appreciable help in convection calculation is the film theory advocated by Langmuir. According to this theory, the heat loss from a surface by convection is conditioned by the pure conduction through a layer of liquid or gas which is at rest (or only in laminar motion), in contact with the

surface. Whilst it is known that such a layer does in fact exist, yet the "films" postulated by Langmuir have no physical existence. The calculated thickness of them is much larger than the actual surface films, though they are of value in calculations because their effective thickness varies in a simple and known manner with the shape and size of the surface, and with the velocity of the fluid.

The papers by Davis[1] and Rice[5] are the earliest in which comprehensive surveys of the available data on convection were undertaken with the aid of similitude theory, the former rejecting and the latter accepting the concepts of the film theory.

Convection is divided into free and forced convection; in the latter, the air motion is produced by means independent of the heating.

As a rough generalization, we may say that in forced convection, the heat loss is proportional to the temperature difference between the hot surface and the mass of fluid in which convection is taking place, whilst in natural convection, the heat loss is proportional to the nth power of the temperature excess. It is commonly stated that n is $\frac{5}{4}$, but whilst this is the value for plane surfaces of moderate size, and for cylinders of considerable diameter, the power may vary on either side of this value for different geometrical systems. In cases where heat transfer is proportional to the area, it may be shown that the value of n is $\frac{4}{3}$.

A generalization which lies nearer the truth is the assertion that in forced convection $H/k\theta$ is a function of Vd/ν, the function being the same for systems of different size but geometrically similar. In these expressions, H is the heat loss per unit of length measured in some consistent direction, V is the linear velocity of the fluid, d a length defining the size of the body, whilst θ is the temperature excess and k, ν are the thermal conductivity and kinematic viscosity of the fluid. These properties vary with temperature, and theory does not indicate at what temperature they are to be taken. It is quite common to take a temperature halfway between those of the ambient fluid and of the surface of the body, but since the heat flow is probably governed by conduction in a narrow boundary layer, it is possible that the value taken ought to be nearer that of the hot surface.

H
V
d, θ
k, ν

The principle of similitude is helpful, but not to the same extent, in natural convection. Here there are three variables instead of only two, but $H/k\theta$ varies only slowly with one of them, so that the results are usually quite manageable when expressed as a family of curves with the latter variable as parameter.

It would be ridiculous in a survey of the present length to enter into details as to the actual results for different geometrical systems, especially since the recent books[15,16] by Fishenden and Saunders and by McAdams deal with the whole literature in great detail. In these books the results of different workers are correlated on similitude principles with considerable care.

§ 7. THERMODYNAMICS

It would not be expected, in a science like thermodynamics, which so nearly approaches the Euclidean ideal of a complete and self-consistent discipline, that much could happen in relation to the foundations of the subject. Nevertheless, just as relativity theory is associated with a disturbance of the Euclidean conception of space, so it has its effect on thermodynamics[7, 13]. The numerical alterations to be expected are small, and apparently only one serious investigator, Tolman[13, 16], has examined the matter. His conclusions are too complicated to summarize here.

In another direction altogether, thermodynamics has suffered a metamorphosis in the last three decades, owing to the incorporation of a third "law," Nernst's heat theorem. For some time it was a matter of doubt precisely what was the scope of the law, but it now appears fairly certain that the vanishing of the entropy at absolute zero is a property only of a crystalline solid and of the ideal gas. In this connexion some remarks in a paper by Giauque and Johnston[6] should not be overlooked.

The nomenclature of the subject has undergone two changes, one beneficial or at least harmless; the other, in the view of the writer, cannot be so described. The first consists in the introduction of the word *enthalpy* to describe the quantity commonly known as total heat or heat content. The other change is the use of the words *free energy* to denote a quantity—thermodynamic potential at constant pressure, or Gibbs's function—other than that which has been designated by these words since the date of Helmholtz's memoir on the *Conservation of Energy*. To make the confusion greater, the symbol F_0 has been used for Gibbs's function, although F has been so widely used for the other function.

Thermodynamics has, ever since its inception, been applied in two quite different directions—in Engineering and in Physical Chemistry. In the former, the position has altered but little in recent years. Even the American trials of a compound heat engine using mercury vapour as operating substance, the exhaust mercury being employed to boil water which operates in other cylinders, are not a new idea. Last century, engines were tried in which the exhaust steam was used to boil ether, though of course the temperature range (and therefore the ideal efficiency) was much lower.

In physical chemistry, thermodynamics has been of continual service, though it has more and more come to be mixed with molecular theories, so that pure thermodynamics tends to recede into the background.

The greatest battles have been fought, and indeed we are still only in a state of truce, over the question of concentrated electrolytic solutions.

The latter subject can hardly be considered as belonging solely to heat, yet it is convenient to consider it here, simply because it falls naturally under thermodynamics, although it is in fact an electrical application.

It will be recalled that, following on Faraday's law of electrolysis, Arrhenius proposed a theory, which won general acceptance, to the effect that polar compounds when dissolved in water (or other solvents of sufficiently high dielectric constant)

became partially ionized. Thus sulphuric acid, for example, produced a certain number of positively charged hydrogen ions and of negatively charged sulphions SO_4 or HSO_4, the former possessing one charge and the latter two.

Thermodynamic reasoning showed that the vapour pressure, freezing point, osmotic pressure and boiling point of such solutions would depend on the concentration of the solution, but measurements of these properties gave the wrong answer for the concentration. The discrepancy was explained in an obvious way by the Arrhenius theory, since the measurements just mentioned did not measure the concentration of the original solute, but the sum of the concentrations of the two products of ionization. Indeed, from a knowledge of these two distinct concentrations, the degree of ionization was easily calculated, and it was considered a complete vindication of the theory when the same result for the degree of ionization was obtained from measurements of the apparently unrelated quality, the electrical conductivity of the solution.

There remained, however, a difficulty, the seriousness of which was not at first fully recognized. The original solute and its products of dissociation are of course in chemical equilibrium, and should thus obey the law of mass action—a thermodynamic requirement which cannot be evaded. This law, applied to the case in hand, imposes a relation between the dissociation and the initial concentration, and is known as Ostwald's dilution law. Observations do not confirm it, the error being sometimes in the ratio of hundreds to one.

The earlier attempts to resolve the difficulty took the form of denying that the concentration itself entered the law, but that some quantity of the nature of an "effective concentration"—called the activity—took its place. In an analogous way, the vapour pressure was replaced by the "fugacity." This method can clearly change the problem into a new one: what are the relations between activity and concentration on the one hand, and between fugacity and concentration on the other? These questions have been the inspiration of an immense amount of observational work of the highest order, but are not yet in sight of answer.

Quite recently, the problem has been attacked with considerable success[1,2] from a new angle by Debye and Hückel. They assume, what had been gradually growing more probable for some time, that ionization, for most substances, is complete or practically so. They thus cut the Gordian knot with respect to Ostwald's dilution law, for if dissociation is complete, its variation with concentration cannot enter into the question. The next step is to show why the measurements of dissociation do not give the true value of unity, and in this the investigators mentioned have had a great measure of success. Their calculation is a detailed one of the motion of an ion in the electric field due to the other ions, when the mass of the moving ion itself is augmented by an envelope of neutral molecules adhering to it. Suffice it to say here that the observed velocities of the positive and negative ions being known, the resulting formula correctly predicts the electrical conductivity. In any case, if the formula is written in the most general form, without any assumption as to the velocities being known or calculable, it shows that the equivalent conductance (i.e. the conductivity/concentration) should be a linear function of the concentration.

The extent to which this theory agrees with experiment has been examined by several observers among whom we mention in particular Ferguson and Vogel[4, 5], and it is certainly a considerable improvement on older theories.

We have already seen that the law of mass action now presents no obstacles, so that the only point remaining is to explain why observations of freezing point depression and analogous formulae give the wrong value for the dissociation. It was asserted above that thermodynamics gave the relation between concentration and freezing point depression, but of course thermodynamic reasoning must give the wrong answer if fed with the wrong data. Here the error lay in assuming laws of too simple a form for the osmotic pressure of the solution. Debye and Hückel's theory gives a value for the osmotic pressure which leads to agreement between the observed and calculated values of vapour pressure lowering and freezing point depression.

Before closing this very brief sketch of the mode of attack adopted by Debye and Hückel, it is needful to remark that although its success has been very great, it must not be assumed complete. The theory purports to deal with more concentrated solutions than the old theory of dilute solutions, but its ideas of strength are modest; $\frac{1}{10}$ normal is about the limit of concentration for which the new equations remain true.

The numerous applications of thermodynamics in physical chemistry are dealt with in books devoted to that subject, such as that edited by H. S. Taylor. We may note, however, that the theory of membrane equilibria, i.e. the relations between two liquids which are separated by a medium impenetrable to at least one component, is not always accurately represented in text-books. The exact position was analysed by Donnan and Guggenheim[10] in 1932, in a valuable paper where the methods used are those originally developed by Gibbs.

§ 8. EQUATIONS OF STATE AND CRITICAL PHENOMENA

The characteristic equations of van der Waals and of Dieterici which accounted for the existence of a critical point, and gave reasonably correct forms to the isothermals in its neighbourhood, marked a stage in the development of the theory of fluids, which we have now passed. No physicist of to-day believes that an equation of state intended to represent accurately the properties of liquid and vapour would be of any use. It would be complicated in the extreme, and for practical use it would be necessary to write it in various approximate forms, according to the region in which it was to be used. For this reason, makers of empirical equations of state (and there are many of them each year) restrict their aim to that of representing either the liquid or the gas, but not both.

Nevertheless the old equations had their uses. Because they contained three constants only, they asserted that there is a "law of corresponding states." Whilst this assertion is untrue, it is not very misleading and is a highly convenient way of estimating properties which have not been measured.

Among the modern equations of state, three are worthy of particular mention. Berthelot's takes the form

$$pv = RT\left[1 + \frac{9\Pi t_1}{128}(1 - 6t_1^2)\right],$$

t_1, T_c, Π where t_1 is the reciprocal of the reduced temperature, and $t_1 = T_c/T$, T_c being the critical temperature, whilst Π is the reduced pressure p/p_c. Thus it accepts the truth of the law of corresponding states. It was derived for pressures from high vacua up to 5 or 6 atmospheres and for temperatures at least 10° or 20° above the critical point. Within these limits it is of remarkable accuracy, and is the one that should be selected where its conditions are fulfilled.

A more ambitious equation is the Beattie-Bridgman equation of state. It contains a greater number of parameters and thus does not admit the validity of any law of corresponding states. The actual equation is

$$pV = RT\left[1 - \frac{c}{VT^3}\right]\left[1 + \frac{B_0}{V}\left(1 - \frac{b}{v}\right)\right] - \frac{A_0}{V}\left(1 - \frac{a}{v}\right),$$

and it is claimed that such an equation represents the properties of many diverse substances with considerable accuracy over a wide range of temperatures and pressures[10].

Callendar's equation of state was put forward to represent the properties of that industrially important and physically rather badly behaved vapour, steam. The equation was developed gradually as the properties were being measured. In its earlier form (as set out for example in Callendar's book on the properties of steam, or in Ewing's *Thermodynamics for Engineers*, it was written in the form

$$p\left(V + \frac{C_0}{T^n} - b\right) = RT,$$

n where n for steam is $\frac{10}{3}$, and b, the "covolume," is of the nature of van der Waals's b b, in that it represents in some way the volume of the molecules. Callendar takes it as equal to the (minimum) volume of liquid water. C_0/T^n is called the "coaggregation volume," a term which exhibits its theoretical explanation.

In the last few years before his death, Callendar extended his total heat measurements to the immediate neighbourhood of the critical point, and came to the conclusion that current theories on that point were not quite correct[11, 12]. According to his observations the presence of even a minute amount of dissolved air vitiates the results. When care is taken to avoid contamination, the meniscus between liquid and vapour disappears at 374° C., but the densities of the two phases are still different. He succeeded in measuring these up to about 380° C., where they were almost identical. Thus if the orthobaric densities of liquid and vapour are plotted on one diagram against temperature, and Mathias's rectilinear diameter drawn, the curves cannot be extrapolated to meet at the observed critical point; the curve usually drawn is too blunt. Actually, beyond the point where the meniscus disappears, there is a rather unstable region where liquid and gas are distinct, the two curves finally meeting at a sharp angle.

The truth of this theory has been but little investigated by physicists, perhaps because it was developed in the course of work for engineers, and was announced in the technical press.

One paper, that by Schröer[8], reports similar phenomena. The experiments were carried out by heating different quantities of the liquid in sealed tubes, above the critical point, and it was found that the densities of the liquid and vapour were not equal until about 0·5° C. above the critical point where the meniscus disappears.

Callendar was led to the above theory of the critical point by adhering to his equation of state, and the related assumptions which he made with regard to specific heat. He confirmed it by the observations just mentioned, and appears to have clinched the matter finally by his direct measurement of the total heat of both phases. These show clearly that there is a latent heat of evaporation above 374° C. and that it only vanishes at about 6·5° C. higher. These total heat measurements also illustrated anew the great influence of even a trace of dissolved air.

To represent the observations in the neighbourhood of the critical point, the equation of state was modified to a small extent; the term b, the covolume, was found to be negative and much smaller than unity (about $-0{\cdot}175$), whilst the coaggregation volume is a more complicated function of temperature than originally assumed. It now takes the form $C_0/T^n\,(1-Z^2)$ where $Z=2{\cdot}39cP/T$, n still having the value $\frac{10}{3}$, and P, T representing pressure and temperature. Callendar regards the factor $1/(1-Z^2)$ as arising from the summation of a geometric series with common ratio Z^2, the physical meaning of the various terms being that there is further complex aggregation among the already aggregated molecules. *P, T*

It is convenient to mention here that Callendar's equation of state is only a part of his theory of the properties of water. Thus throughout his work he makes the assumption that, when water and its vapour are in equilibrium, the former contains its own volume of "steam molecules" dissolved in it. For this reason the specific heat of the water is taken as the sum of two terms, one due to the "water molecules" and one due to the necessity of evaporating a certain amount with each rise of temperature. The resulting equation is simple, and represents the results well above the minimum of the specific heat curve.

Remembering that Callendar's equations are based entirely on total heat measurements, and on the experimental adiabatic law $P/T^{13/3}=\text{const.}$, it is interesting to note that they agree with direct measurements, due to Knoblauch, on the specific volume of steam at pressures up to saturation, in the temperature range from 0° to 180° C.; moreover, if the vapour pressure of steam is calculated from the equation, it agrees with the measurements of Holborn and his colleagues all the way from 0° to 374° C., and it automatically ceases to predict a vapour pressure at all beyond 380·5° C., the point where experiment shows that the latent heat vanishes.

Leaving now the region of the critical point, and passing to the neighbourhood of the freezing point of a liquid, a phenomenon which had been overlooked until recent years has to be chronicled. This is the existence, in several cases, of a second liquid form, stable immediately above the freezing point and differing from the

liquid form stable at higher temperatures. The case which has been studied in greatest detail is that of helium, where the boundary curve between the two forms on a temperature-pressure diagram has been delineated. The two varieties of liquid differ in their dielectric constants, specific heat, density and latent heat of evaporation.

Other liquids which show the effect have been investigated by Wolfke and Mazur[14], and it has been suggested that it may be general. The phenomenon recalls the cybotatic theory of liquids, which claims that in the fusion of a crystal the atoms do not at once assume the completely random arrangement usually accepted as the true model of a liquid, but that traces of the orderliness of the crystal structure remain over a certain temperature range. This theory, of course, would not account for the *sudden* changes observed in the cases of helium, ether, etc.

Both the cybotatic theory, and the observation of the phenomenon just described, call to mind the peculiar behaviour of water, with its maximum density and minimum specific heat, but neither can explain it so well as the theory that water consists of a mixture of polymers of the simplest molecule, H_2O.

The case of helium was referred to in an earlier section of this report, when dealing with the attainment of low temperatures, and literature references are given there in respect of the fact that there is no triple point between the vapour and the two liquid forms.

§ 9. VAPOURS

In the eighties of last century Rankine integrated the Clapeyron equation approximately, and thereby reached a formula for the saturation pressure of any vapour, as a function of temperature, his result being $\log p = A - B/T - C/T^2$.

Almost all subsequent efforts to obtain a more accurate formula for this relationship have consisted in the empirical addition of terms to the Rankine equation, such additions being justified by the fact that Rankine had to ignore the variation of latent heat with temperature, neglect the volume of the liquid in comparison with that of the vapour, and assume that the latter obeyed the law of an ideal gas.

If the additional terms are positive or negative powers of the temperature, it is evident that the equation can take no cognizance of the existence of the critical point. This defect can be remedied by the inclusion of a term depending on $\sqrt{(T_c - T)}$ or $\log (T_c - T)$, where T_c is the critical temperature, since either of these expressions becomes imaginary as soon as T exceeds T_c; if the logarithmic form is used it must occur in the denominator, since it is infinite at the critical point.

An example of the complexity required in a vapour pressure equation if it is accurately to represent good data may be seen in that given in International Critical Tables, volume 3, to represent the properties of ammonia.

From $-78°$ to $+70°$ C.,

$$\log_{10} P = a - b/T - c \log T + dT + ct^2,$$

and from $+50°$ to $+132{\cdot}9°$ C.,

$$\log_{10} P = A - \frac{T_c - T}{T}[B - C(T_c - T) + D(T_c - T)^2 - E(T_c - T)^3],$$

where P = vapour pressure, T = temperature, T_c = critical temperature and the remaining letters are constants, which are given to six or more figures.

An alternative method of obtaining a rational vapour pressure equation would be from the equation of state, assuming Maxwell's theorem of the equality of area on a p, v diagram between the points where the horizontal line representing the vapour pressure cuts the corresponding isothermal for the first and second, and for the second and third times.

The only example of a procedure resembling this appears to be that of Callendar, who calculated the vapour pressure of water in this way from his equation of state, obtaining values in close agreement with experiment. It is to be remembered, however, that in his case the equation of state consists of two quite distinct formulae, one for the liquid and the other for the vapour, and that the calculation is not formally on the lines sketched above. It proceeds by equating Gibbs's function for the two phases.

The experimental technique of vapour-pressure measurements has altered considerably in recent years, though modern methods are still based essentially on the ideas of static or dynamic measurements. The alterations are not primarily in the technique of heat measurements, but in that of general mechanics.

At one time, the only means available for the measurement of high pressures with accuracy was to construct a tall mercury manometer. Amagat, for example, worked at the bottom of a mine so that his mercury column could extend up the shaft. The obvious alternative of loading a piston with weights had to await developments in the machine shop, for at high pressures the piston must not permit leakage, and yet the readings will be erroneous if there is any friction on it. These conflicting requirements can now be met by the accurate grinding of the surfaces of the piston and of the tube in which it slides, aided by the device of keeping the piston in continuous slow rotation.

At the other extreme it has now become possible to measure with accuracy pressures lower than a liquid manometer could even detect. This is a by-product of the immense amount of high-vacuum work carried out in the last three decades, but it must be remembered that not all the gauges devised in that field are useful here. The McLeod gauge, and several others, ignore the pressure of any vapour which may be present, and are therefore useless for the present purpose.

An interesting method of measuring the very low pressure of mercury vapour in the neighbourhood of ordinary room temperature was used by Knudsen. In this method, the law connecting the rate of flow of the vapour through an orifice with the pressure difference on the two sides of the orifice is assumed to be known. The pressure on one side of the orifice is the vapour pressure itself, the pressure on the other being made negligibly small in comparison with it. The rate of flow is deduced by collecting and weighing the mercury which passes through the orifice.

A somewhat similar method has been used by Mayer[5] and by Egerton[4]; the former measured the rate of flow by the deflection of a quartz fibre projecting across the orifice.

Closely connected with the subject of vapour pressure is that of condensation of vapours. If a vapour is gradually cooled, then as soon as the temperature reaches that at which the applied pressure would be the saturation pressure, the equilibrium state is one in which condensation takes place. This condensation is frequently delayed, particularly if the atmosphere is dust-free, when in fact eight-fold supersaturation seems to be the limiting point. This effect was at one time believed to be connected with the increase of vapour pressure over a curved surface, but it is now thought that the case is one of true metastability. There is no doubt of the influence of nuclei, but it is almost certain that these are hygroscopic, and exert their influence on this account, and not primarily because of their size.

When negative electrically charged particles are present, the supersaturation can only reach about two-fold, and the use made of this fact by C. T. R. Wilson in his expansion chamber is very familiar to every physicist.

Condensation and evaporation are processes of great technical importance in chemical engineering, and also in mechanical engineering; in the former we have distillation on the large scale, and in the latter the boiler and condenser of a steam engine exist solely as vessels in which these phenomena can occur.

Until recently the attitude had been universally adopted, that "condensation is condensation" and "evaporation is evaporation," but now it is being realized that even in such simple processes there is much to learn. Thus a number of German investigators have recently studied the process of condensation of saturated and of superheated vapour in metal tubes. If the tubes are vertical they find that two regimes can occur, in one of which the condensed water forms a continuous film on the surface, and in the other the condensate is present in the form of drops. The heat transfer in the two cases is vastly different, the drop regime being superior in a ratio which may be as high as 6 to 1.

The explanation of this difference is that in the film regime the vapour condenses and gives up its latent heat to the inside of the film, through which the energy is forced to travel by conduction, whereas in the other the drops form directly on the metal, and there is no obstacle to the flow of the heat liberated.

In a somewhat similar way, the bubbles formed in the process of evaporation vary greatly in size, with a corresponding alteration in the efficiency of the evaporation process. It had long been known that oil in the feed water to a boiler could lead to burning out of the water tubes, and the explanation of this is implicit in the results of this work. If there is a film of oil on the metal surface, evaporation takes place over large areas of the surface, the liquid appearing to swell up into large bubbles. This condition is the analogue of the film regime in condensation, and is accompanied by relatively poor heat transmission; but in evaporation problems the rate of supply of heat is fixed independently of the process, so that to deal with this supply the temperature differences become large and so the metal is overheated.

Quantitative measurements have confirmed the general theory, and in particular

the rate of increase in volume and surface of the bubbles has been studied by measurement of high-speed cinematograph pictures.

A very interesting survey of the recent work has been given by Jakob[18].

An entirely different aspect of the evaporation problem is that relating to the process when it occurs at normal temperatures, and at normal or low pressure. The laws governing evaporation under these conditions are clearly fundamental in the theory of the wet and dry bulb hygrometer, and must also be of interest to the meteorologist, since evaporation on the large scale clearly affects climate.

The problem is also of interest to the molecular physicist. If the pressure over the evaporating liquid is known, the number of molecules impinging per second on the surface can be calculated by kinetic-theory methods, and the number leaving the surface in equilibrium is equal to this number. When evaporation is occurring, the number leaving the liquid must be the same as if equilibrium had been established, but the number impinging is less. Now the measured rate of evaporation shows the difference between the number impinging and the number leaving. The actual results are interpreted as showing the probability of reflection for a molecule striking the surface. Much of the work on this part of the subject has been carried out by Alty[20], and the device used in one of his researches to measure the actual surface temperature of the liquid may be of use to workers in other fields. The difficulty in these cases is that the evaporation itself leads to cooling, so that the surface temperature which determines the evaporation is of necessity lower than that measured at any finite distance in the liquid below the surface. Alty overcame this by measuring the surface tension (which also depends on the ultimate layer) and using this to deduce the temperature.

From the point of view of practical applications, the researches on the rate of evaporation under given conditions are noteworthy mainly for their discordance.

It would be expected that the rate of evaporation would be proportional to the difference between the saturation vapour pressure and the pressure of the water vapour in the ambient air. Although most workers confirm this, others have found that it does not describe the facts exactly.

When the air above the liquid surface is in motion, the evaporation naturally increases, but the precise shape of the (velocity, evaporation) curve is in very considerable doubt.

It is also to be noted that evaporation is not proportional to the area of exposed surface, the law apparently varying from the square root of the area for small surfaces, up to proportionality in the case of lakes and other large surfaces.

Doubtless much of the discrepancy which is found in the results of different workers is to be sought in differences in the turbulence of the air which passes over the surface, this turbulence being in many instances caused by the apparatus itself, since a vessel containing water cannot easily be streamlined to the approaching air. In an investigation now being carried out at the National Physical Laboratory by R. W. Powell and E. Griffiths[24], the water is exposed only on a layer of muslin cloth, and streamlining is thus possible.

Finally, in connexion with the detailed mechanism of evaporation and conden-

sation, we must notice the work of Langmuir and Knudsen, which is dealt with later in the section on kinetic theory.

§ 10. HYGROMETRY

Gas analysis is a branch of chemistry, and not of physics, but there is one case, that of determining the ratio of water to the other constituent in a mixture of water and another gas, which is generally regarded as falling within the province of physics. The primary reason is that the methods customarily used for this determination are of a physical nature, though, indeed, a secondary reason is that the results of the measurement are as often needed in physics (which may be taken to include meteorology) as in chemistry.

Apart from chemical analysis by absorbing and weighing the water, the standard methods for hygrometry are all physical. They are the dew point method, the wet and dry bulb method, and the extension of hair method.

The first of these has never been improved beyond the point at which Regnault left it; after his time, and until quite recently, common practice had fallen behind the point reached by him, in that it was usual to ruin the thermal contact between the silver thimble and the glass tube by a layer of a cement. This fault is not now committed.

The theory of the wet and dry bulb method is only now beginning to be really well understood, though experiment had earlier shown the importance of an air current over the wet bulb. This can be attained in various ways, the best known being by the use of a clockwork fan, as in the Assmann instrument, or by whirling the instrument round as in the sling psychrometer. Instruments of both these forms are obtainable commercially. In addition to the development of the ventilated psychrometer using mercury thermometers, a number of experimenters have produced instruments using either thermocouples or resistance thermometers for measuring the temperatures. Some of these instruments have been adapted for reading at a distance, and another feature in some cases has been the measurement, not of the wet-bulb and dry-bulb temperatures separately, but of the dry-bulb and wet-bulb depressions. This permits of the latter being known more accurately and is especially desirable at low temperatures; full advantage cannot be taken of the instrumental accuracy, however, until better tables are available below 0° C.

The hair hygrometer is hardly a physical instrument in the strict sense, but its utility is undeniable. It has in recent years been subjected to careful study[2] from the point of view of the permanence of its calibration by Dr E. Griffiths.

In addition to the three methods which have been designated above as standard methods, at least a dozen others have been proposed and have met with more or less success. Methods based on the change in electrical resistance of some material have been particularly popular, among the substances tried being graphite, cotton-wool, glass-wool and a roughened glass surface.

The heat evolved when water is absorbed by cotton wool has also been suggested as a means of measuring humidity, as has also the change in refractive index

of a glycerine solution as its concentration alters. Numerous alternatives to the human hair have been tried, noteworthy among them being horse's hair and gold-beater's skin.

Distinct from the problem of measuring humidity, is that of controlling it. This is a problem of applied physics with great industrial importance. The favourite methods are (1) to saturate air at one temperature (thus determining its moisture content once for all) and then to warm it to some other temperature, the one desired, or (2) to use a hygrometer of the hair type, allowing the extension of the hair to operate a mechanism controlling the supply of moisture to the enclosure or air stream concerned.

§ 11. KINETIC THEORY OF GASES

In the great upheaval attending the birth and early youth of the quantum theory, the old kinetic theory was badly battered from time to time, and in particular its theorem of equipartition of energy suffered severe attacks. This theorem had always led a hard life, no mathematical physicist being satisfied that anyone but he himself had yet succeeded in giving a rigorous proof of it. It had become accustomed to this, but the attack was now of a different nature—for the first time its truth was in question.

The problem of equipartition was of course far wider than its simple dynamical application to gas theory, and not only is its untruth in certain fields now well understood, but the correct laws to replace it are known. In the theory of gases no alteration is made to the old doctrines, save at exceedingly low temperatures where degeneration sets in. This phenomenon was first discovered experimentally when Eucken[3] found that at low temperatures the specific heat of hydrogen fell first to the value usually associated with a monatomic gas, and then (below 20° K.) to still lower values. Later work has only tended to confirm the belief that the specific heat practically vanishes altogether at sufficiently low temperatures, a result which was applied by Sommerfeld to explain why the electron gas in a metal does not contribute to its specific heat, although it is responsible for electrical and thermal conduction. The theoretical explanation of these experimental facts followed quickly on their announcement.

Pauli's Exclusion Principle had been generally accepted; it stated that no two electrons in an atom could have all their quantum numbers the same, i.e. could be indistinguishable. It was only necessary to assume quantization of the translatory motion of molecules and to extend the principle to the molecules as a whole to make the explanation obvious. At low temperatures, where the average energy is small, each molecule assumes the lowest unoccupied energy level and cannot budge when the temperature is further reduced, since the lowest energy levels are all occupied.

Since pressure is a result of the bombardment of the walls of the containing vessel by the molecules, it is to be expected that pressures will not vary with temperature when the degenerate state has been reached, and this has been verified.

Mention of the kinetic theory naturally calls to mind the name of Knudsen, who has devoted some thirty years to a theoretical and experimental study of its conclusions, particularly in the realm of low pressures, where the mean free path of the molecules is large compared with the dimensions of the vessel.

In such vacua, the law of Poiseuille for viscous flow through a tube no longer holds. The actual law is simpler in form, and in fact the flow in grams per second is $k\sqrt{\rho}.(p_2-p_1)\,S^2/lC$, where ρ is the density of the gas at some standard (fixed) pressure, and (p_2-p_1) is the pressure difference between the ends of a tube of length l, circumference C and cross-section S. Gas flow follows this law from the lowest attainable pressures up to a certain value (e.g. about 0·02 mm. of mercury in the case of CO_2), after which there is a region not yet explored theoretically, before the Poiseuille type of flow sets in. Physically the difference between the Knudsen and the Poiseuille types is one of degree. In both cases (and presumably also in the intermediate region) the retarding forces on a layer of gas are due partly to the influence of the walls and partly to adjacent layers; at ordinary pressures the wall effect is completely negligible, whilst in the Knudsen region it is predominant.

k, ρ
p_2, p_1
l, C, S

Knudsen's calculations yield not only the form of this law, but also the numerical value of k, the results having been experimentally confirmed with a high degree of accuracy. The value of k would vary with the distribution law assumed for molecular velocities, and hence the experiments provide a test of the truth of the Maxwellian law.

More direct proofs[8,9] have been given by experiments in which a stream of molecules obtained by vaporization is forced by slits to travel in a straight line. In the path are two discs rotating at high speed, each with one or more apertures. Evidently the molecules which pass both discs are travelling with such a speed that they traverse the space between the discs in the same time that the slits take to travel round to the appropriate position. Thus a velocity analysis is possible. The results confirm Maxwell's law within experimental error.

A calculation for the rate of flow of a gas through a small aperture when the pressures are different on its two sides was also carried out and experimentally checked by Knudsen, who then proceeded to make an ingenious application of it, already referred to in the paragraph on the measurement of vapour pressures. For example, the vapour pressure of mercury at 27° C. is about 0·002 mm. of mercury, a quantity which was measured with an accuracy of 1 or 2 per cent by the method referred to. The apparatus, which is exhausted, consists of two vessels in communication via a small orifice, the mercury occupying the upper one. This chamber is maintained at a fixed temperature between 0° and 30° C., so that it becomes filled with mercury vapour at the saturation pressure. The lower vessel is then surrounded with solid carbon dioxide so as to reduce the vapour pressure in this part of the apparatus to an extremely low value. The rate of flow of vapour through the orifice between the two is deduced by weighing the quantity of mercury which is collected in the lower part of the apparatus after a known time. The mean flow is thus known, under a pressure difference equal to the saturation pressure of

mercury in the upper half of the apparatus, and consequently this pressure itself can be calculated.

Important as this work is in providing a technique for measurements of such low vapour pressures, it leads up to a still more important matter when attention is focused on the lower part of the vessel. If this is surrounded by solid carbon dioxide, the mercury adheres to the glass walls of the vessel, forming a mirror there. It appears in fact that a molecule striking the walls gives up its energy completely. If a molecule strikes a wall maintained at a somewhat higher temperature it is more likely to be reflected. The change from one state of affairs to the other is not gradual, but quite abrupt. The effect was discovered independently in the same year by Knudsen himself, and by Ward[4a]. The temperature at which reflection ceases, and all the molecules remain at rest on the surface, is a characteristic temperature for each vapour. The effect was at first doubted by Langmuir and others, but an examination of the photographs given by Choriton and Semenoff[5] leaves no doubt of its reality.

Another question which arises in connexion with the behaviour of molecules which strike a surface is the velocity with which they return. In general, this is neither the temperature with which they hit the surface, nor that of the latter itself.

The discussion of such questions is made precise by introducing the conception of an accommodation coefficient. Numerically, this is defined as the ratio $(T_1-T_2)/(T_1-T_0)$, where T_1 is the temperature (defined by the energy of the molecule) of a molecule striking a surface at temperature T_0, and T_2 is the temperature of the same molecule after reflection at the surface. It may be remarked that there is no reason beforehand to assume that the accommodation coefficient for translational energy is equal to that for the rotational and vibrational energy. This depends on whether the effective temperature is the same for both energies, a question which must be answered by experiment. Knudsen finds, as a matter of fact, that for diatomic gases these quantities are equal. In studying such impacts at surfaces, a fundamental law is the cosine law, according to which the molecules leaving a given element of surface in a given time are so distributed that the number leaving at an angle θ to the normal is proportional to $\cos\theta$. Such a law is suggested by a consideration of an equilibrium case, where the gas and surface are at the same temperature, so that the number of molecules approaching the surface in a given direction must on the average equal the number leaving in that direction. It is easily shown that the impinging molecules follow this law, and consequently in equilibrium the departing molecules must do so. Knudsen proceeded to make the assumption that this is so, even when there is no equilibrium.

T_1
T_0, T_2

On the basis of this law, some important results follow. It is not difficult, for example, to show[7] that in a narrow gap between two surfaces which are maintained at θ_1, and immersed in a gas at θ_2, the pressure between the two surfaces is higher than in the surrounding gas, in such a way that $p_1/\sqrt{\theta_1}=p_2/\sqrt{\theta_2}$. This leads to a consideration of the force tending to separate two surfaces lying close together, and maintained at different temperatures θ_1 and θ_2. The force per unit area is

θ_1, θ_2

p $\frac{1}{2}p/\{\sqrt{(\theta_2/\theta_1)}-1\}$, where p is the pressure of the mass of the gas in which they are immersed. This force is one which accounts for a part of the radiometer action met with in high vacua.

The same force has been applied by Knudsen in the construction of his absolute manometer, one of the few instruments which can measure a very low pressure (including that of any vapour present) without needing calibration. In one form of this there are two plates, one of which is heated electrically, and the force on the movable one is measured by means of a torsion balance. The temperatures are also measured and hence p in the above equation can be calculated.

Another consequence of the cosine law, combined with the notion of an accommodation coefficient, is that the simple theory of conduction in gases given by Maxwell has to be modified. This effect has been discussed in a long paper[13] by Knudsen, and has been the basis of a good deal of the experimental work on accommodation coefficients.

The theory of the accommodation coefficient on wave mechanics has been attacked by Mott[14], and some of the best experimental work is that of Roberts, who finds that purity of a metal surface affects its accommodation coefficient profoundly.

§ 12. THEORY OF CONDUCTION

The law of Wiedemann and Franz, which stated that the ratio of thermal to electrical conductivity was the same for all pure metals, was both extended and explained by the electron theory, as soon as the assumption was made that the valence electrons of a metal could lead a free existence, detached from any individual atom. The extension consisted in a recognition of the fact that this ratio has a temperature coefficient, being in fact proportional to the absolute temperature (a prediction of the theory confirmed by experiment). The theory, which was largely developed by Drude, led to a difficulty of a different kind, however. If the electrons are free in a metal, and if thermal conduction is due to the transport of energy by the passage of those electrons with high velocity, then they must contribute to the specific heat of the metal; otherwise those in the hotter region would not possess the extra energy which in fact they are observed to carry to the cooler parts. Detailed calculation showed that the contribution to the specific heat by the free electrons was practically the whole quantity, implying that the rest of the structure contributed nothing. This paradox was the more serious because the high coefficient of expansion of metals shows that energy does go into the lattice, and also because non-conductors, such as sulphur, do not have abnormally low specific heats.

The difficulty remained unresolved for practically twenty years, until Sommerfeld[3] suggested that the electron gas is in a state of degeneration. We have already seen that in such a state experiment as well as theory shows that the specific heat becomes small—even vanishingly so at sufficiently low temperature. Fortunately the suggestion has not to remain a mere assumption, introduced to provide a way out of the impasse. The theory yields a connexion between the molecular weight of a gas and the temperature at which it will show any given degree of degeneration,

from which it is easy to prove that at any temperature up to 36,000° C. electron gas in a metal must be as degenerate as helium at 5° K., a temperature where its specific heat is very small. It must be noted, however, that Sommerfeld's theory is incomplete, in that it takes no account of the interaction of the electrons with the positive ions forming the crystal lattice of the metal. When this is allowed for, the results are slightly (but not qualitatively) altered.

We may thus conclude that on the theoretical side the electronic theory of metals is well established. In the paper to which reference has been made Sommerfeld also showed that, in its new form, the theory is fairly successful in dealing with the Peltier and other thermo-electric effects, so that it may certainly be regarded as unlikely to undergo substantial alterations in the future.

On the experimental side it is found that the main prediction of the theory is substantially true, but it is by no means exact. The Lorentz function $K/\lambda T$, where λ is electrical and K is thermal conductivity at absolute temperature T, ought to be a constant, independent of temperature, for all metals. Neither condition is strictly fulfilled, and various theories have been put forward to account for this, though none seems fully convincing. The variation may be due to either λ or K or both failing to follow the predicted law. One suggestion which has been put forward in several quarters must certainly contain some truth. It is pointed out that non-conductors of electricity are not devoid of thermal conductivity, so that it is plausible to assume that in a metal there is an additional mechanism for thermal conduction, similar to that operating in non-metals. The most detailed, and at the same time the most successful, treatment on these lines is that of Bidwell[5]. He finds that the ratio of thermal conductivity to specific heat should be of the form $B+A/T$, where T is absolute temperature, and is able to confirm this formula by experimental results. This double mechanism would mean that the Lorentz function, as deduced from experimental results, would be $(K+K_0)/\lambda T$, where K_0 varies from metal to metal, and does not obey the same law of temperature variation as K. Another point which has to be remembered is the effect of impurities on the electrical resistance. Quite a small admixture of another substance will usually raise electrical resistivity appreciably; in addition, if the alloy is not completely homogeneous, then thermo-electric effects at the junction between the components will operate to raise the apparent resistance, as pointed out by the late Lord Rayleigh.

λ, K, T

K_0

Although the electronic theory is of great value in the case of metals, its main success has been in the prediction of the constancy of the Lorentz function. In the theoretical deduction of this relation, most of the constants disappear from the result, so that it can be immediately tested. The theory is not so successful in dealing separately with the electrical and thermal conductivities, and in particular the quadratic law on which platinum resistance thermometry is based is not a consequence of the theory. The variation of electrical conductivity with temperature (a subject outside the present section of the report) is correctly predicted, but only for substances which do not become supra-conducting, and no prediction of value is made as to the variation of thermal conductivity with temperature.

When we turn to the subject of non-metals, theory gives no assistance whatever.

It is found for most materials that conductivity increases with increasing temperature, thus adding to the difficulties of insulation at high temperature. A parabolic formula usually suffices to represent the results sufficiently well, but this is entirely empirical, and in a certain case[6] the formula $K = A + B \log T$ is found to give better results, where K is conductivity and T is absolute temperature. It is to be noted that no observations at very low temperature appear to have been made; it is conceivable that determinations on a few materials of definite composition, say elements, oxides (including silica) and chlorides, down to liquid hydrogen temperatures, might bring to light some interesting facts. They would at any rate show the general form of the typical curve, of which at present we only know the upper portion.

Among the non-metals is carbon, which is interesting in being a fairly good conductor both of heat and of electricity. Recent observations have shown that for it the Lorentz function is abnormally large. It decreases rather rapidly with temperature, approaching the value for a metal, though not reaching it even at 1200° C. This seems to suggest that in carbon there is electronic conduction, but that the additional non-electronic part is larger than for a metal, and that this additional part diminishes relatively rapidly as the temperature is increased. The effect is found not only in amorphous carbon, but also in graphite, and the only surprising thing about it is the definite evidence which the phenomenon offers that the carbon atom can so easily lose a single negative electron.

§ 13. THE SPECIFIC HEATS OF SOLIDS

As we have seen in considering thermal conduction in metals, the free electrons in them contribute nothing to the specific heat, so that the mechanism which takes up energy when the substance is warmed must be the same for metals as for non-metals.

The laws of variation of specific heat with temperature are so much simpler for crystalline solids than for liquids, that we are bound to seek an explanation in the regularity of the space lattice of the solid, in contrast to the random arrangement of the molecules in the liquid.

In the early days of the quantum theory Einstein showed that, by assuming the presence of a number of oscillators with quantized energy, the general form of the specific heat curve, and in particular the vanishing of this quantity at low temperature, could be explained.

A step in advance was made by Debye, who guillotined the spectrum arbitrarily at a fixed high frequency, on the ground that the substance could not have an infinite number of degrees of freedom if made of discrete molecules. It was from that time clear that the true theory must take account of the crystal structure of the substance, though the difficulty of the mathematical analysis deterred many from attacking the problem. Two very eminent mathematical physicists, Max Born and von Kármán, have given the solution[1] for the simplest type of crystal lattice, but

their result, shown below, is almost too formidable for numerical treatment. Their formula is

$$C_v = \frac{Vk}{\Delta}\left\{\sum_{i=1}^{3}\left[4D\left(\frac{\theta i}{T}\right) - 3P\left(\frac{\theta i}{T}\right)\right] + \sum_{i=4}^{3p} S\left(\frac{\theta i}{T}\right)\right\},$$

where V is the atomic volume, k is Boltzmann's constant, Δ the density, θ the characteristic temperature and T the actual temperature. p is the number of atoms in the unit cell whilst θ may of course be replaced by $h\nu/k$, if desired, where h is Planck's constant, ν the frequency, and k the gas-constant for one molecule (Boltzmann's constant). The frequencies are obtained, the one set from mechanical, and the other from optical, data. The fact that the summation for the first two terms is different from that for the last term, corresponds to the fact that they are the natural mechanical frequencies, whilst the others are the optical (infra-red) frequencies, which are exhibited most clearly by the well-known method of residual rays. D, P and S are abbreviations for the functions $\frac{3}{x^3}\int_0^x \frac{x^3\,dx}{e^x-1}$, $\frac{x}{e^x-1}$ and $\frac{x^2e^x}{(e^x-1)^2}$, respectively, so that the square bracket represents a Debye function, and S is an Einstein function.

V, k, Δ, θ
T, p
θ
ν
D, P, S

The basis of the above theory is the assumption that the vibratory energy of the elements of the crystal lattice is quantized, and that the crystal is arranged in the familiar cubic form. It is assumed that the neighbouring atoms act on each other with a force which varies as a simple power of the distance, and it is easy then to see that the crystal will have certain natural frequencies, just as a string loaded with masses at certain points would do. Moreover, there is no need for an *arbitrary* cut-off at a fixed high frequency, since the finite separation of the lattice-points effectually sets a lower limit to the wave-length, and hence an upper limit to the frequency. The quantization laws, even on the older quantum theory, presented no difficulty, the underlying theory being quite straightforward, but the mathematics is very complex, and for crystals of other classes the theory has not yet been worked out.

It should be remarked that the specific heat measured experimentally for a solid is always that at constant pressure, though in the first place the theory gives that at constant volume. The difference could easily be calculated from a knowledge of the elastic constants, but these are not usually known at very low temperatures, where the correction is most important and where the form of the specific heat curve is of most interest. Several methods of evading the difficulty are known, the soundest and most instructive being based on Grüneisen's law. This law states that the ratio of specific heat to coefficient of thermal expansion is independent of temperature. If this ratio is determined at the lowest temperature for which all the data are available, it can be extrapolated to lower temperature with confidence, since Grüneisen's law has both theoretical and experimental support. In particular, the lattice theory of Born leads to it, as shown in his paper[3] with Brody.

It may also be noted that in later papers Born has accounted for the rise of specific heat beyond the value $3R$ which sets in at higher temperatures. He finds that the rising part of the curve, beyond this equilibrium value, should be asymptotic to a straight line passing through zero specific heat at the absolute zero of temperature,

though the experimental evidence in support of this deduction is not decisive. It has been considered in particular by Magnus.

§ 14. THE SPECIFIC HEATS OF GASES

When a gas is heated, the energy received goes to increase not only the translational velocity but also the angular velocity and the internal vibration of the molecules. The energy corresponding to each of these motions is probably quantized. In addition we know from experiment that the molecules can be raised to energy states other than the ground state—in less general terms, the outer electrons can be raised to other energy levels or orbits—so it is to be expected that some energy will be disposed of in this way.

At sufficiently low temperatures, the quantum for any of the forms of energy is so much greater than the average energy of a molecule, that very few possess more than the minimum demanded by the exclusion principle. The gas is then degenerate and its specific heat is zero. At high temperatures, the translational energy begins to increase, and if this were the only factor operative, the gas would ultimately have molecular heats of $3R$ and $5R$ at constant volume and pressure respectively. The paper by Eucken in which this behaviour was first reported experimentally has been mentioned in the section on kinetic theory.

When the temperature is raised still higher, the average energy is large enough for molecules with a quantum of vibrational or rotational energy to be statistically important. We should at this stage expect the specific heat-temperature curve to rise again, the whole relation being the sum of a number of functions each relating to one form of energy. At still higher temperatures, the electronic structure would absorb energy.

Now there is a totally different way in which the energy other than translational can be studied. The band spectrum of a molecule is due to transitions of the molecule from one energy state to another. In general, the quantum numbers for rotation, for vibration and for electron configuration all change in the production of any one line, but by the labours of the spectroscopists it has become possible to identify the lines due to a change in any one energy form alone. Moreover, from the spectroscopic data, the moment of inertia of the molecule (which must alter when an electron changes its orbit) can be calculated. Thus the energy in the different states (reckoning from the lowest as an arbitrary zero) can be precisely known from spectroscopic theory. From this, it is not a difficult matter to deduce the form of the theoretical relation between temperature and heat capacity for these forms of energy. If the translational specific heat is then added to the sum of the others, the theoretical specific heat can be compared with observations.

This has been carried out by various investigators for different gases, and the results on the whole confirm the theory. The earlier papers dealing with the rotational and vibrational energy of gas molecules were written long before the systematic formulation of present-day quantum mechanics. Consequently they take the Bohr atom as their starting point. It is curious that the first calculations

of the contributions of these energies to the specific heat were all made *ab initio* from the properties of a model atom, and not from spectral data. An interesting paper of this type is the one by Kemble and van Vleck[2], in which the energy of an atom which expands on rotation owing to centrifugal force is calculated.

The first paper in which the band-spectroscopic data were utilized appears to be that by Hicks and Mitchell[4]. They took the energy of successive states direct from the wave numbers of the measured bands of hydrogen chloride (after the assignment of the appropriate quantum numbers). The only problem is then the calculation of the probability of each of these states at any given temperature. For this calculation, they assumed the older form of law and their calculation was later corrected by Hutchisson[6]. His curve shows a pronounced maximum in the specific heat at about 12·5° K., a temperature at which there is no experimental check available. For the one gas in which observations are extensive enough to show such a maximum if it existed, hydrogen, no maximum is predicted. This is not because the curve has none, but because ordinary hydrogen is a mixture of ortho- and para-hydrogen and the observed curve would be an average. The existence of these two forms was first suggested by Heisenberg, and an account of the experimental facts will be found in the paper by Giauque and Johnston[7]. All the earlier work was concerned with diatomic molecules, and it was not until 1932 that any papers appeared dealing with triatomic gases. This field is now being opened up with considerable success, and the theory for such gases as CO_2 should soon be complete. This gas and steam (the most important substance industrially) are of course still being studied experimentally[1], the results of throttling (Joule-Thomson porous plug) experiments being of considerable use in the preparation of tables.

It may be noted that there is no evidence that energy is distributed instantaneously among the various possible forms, so that conceivably the well-known velocity of sound method for determining the ratio of specific heats may be seriously in error if such a high frequency is used that the wave travels faster than the molecules can accommodate themselves to the equilibrium position of energy. Such a possibility has been studied by Kneser[10] among others.

On the experimental side, the work up to about 1923 has been ably summarized in a book by Partington and Shilling[3], who deal also with the calculation of the correction from constant volume to constant pressure specific heats. No radical alterations in technique have occurred, but it may be noted that in the velocity of sound method it is now possible to use high frequencies, thanks to the quartz crystal oscillator[16].

REFERENCES

Section 1

Descriptions of low temperature equipment.

(1) W. Meissner. *Phys. Z.* **26**, 689 (1925).
(2) P. Kapitza and J. D. Cockcroft. *Nature*, Lond., **129**, 224 (1932).
(3) W. Meissner. *Z. Ver. dtsch. Ing.* **76**, 580 (1932).
(4) F. A. Lindemann and T. C. Keeley. *Nature*, Lond., **131**, 191 (1933).

Miscellaneous.

(5) T. H. Laby and E. O. Hercus. *Philos. Trans.* **227**, 63 (1927). (Determination of *J*.)
(6) E. Griffiths. *Nature*, Lond., **128**, 485 (1931). (Definition of calorie.)
(7) W. J. de Haas. *Nature*, Lond., **132**, 372 (1933). (Lowest temperature reached.)

Section 2

On thermometric scales and fixed points.

(1) F. P. G. A. J. van Agt and H. K. Onnes. *Proc. K. Akad. wet. Amst.* **28**, 693 (1925). (Comparison of gas thermometers.)
(2) *Bur. Stand. J. Res.*, Wash., **6**, 1119 (1931). (Melting point of Pt.)
(3) W. E. Forsythe. *Phys. Rev.* **38**, 1247 (1931). (International comparison.)
(4) *Rep. nat. phys. Lab.* **29** (1933). (Melting point of Pt.)
(5) F. W. Roeser, F. H. Schofield and H. A. Moser. *Ann. Phys.*, Lpz., **17**, 243 (1933); *Bur. Stand. J. Res.*, Wash., **11**, 1 (1933).

On instruments for thermometry.

(*a*) *Thermostats.*

(6) S. J. Lewis and F. M. Wood. *Trans. Faraday Soc.* **17**, 696 (1922).
(7) E. Griffiths. *Methods of Measuring Temperature* (London, 1925).
(8) J. A. Beattie. *Rev. sci. Instrum.* **2**, 458 (1931).
(9) P. Chevenard. *Rev. Métall.* **28**, 453 (1931).
(10) R. B. Sosman. *Industr. Engng Chem.* **23**, 1369 (1931).

(*b*) *Thermometers and pyrometers.*

(11) Anon. *Engineering*, **110**, 736 (1920). (Mercury in steel thermometer.)
(12) W. E. Forsythe. *J. opt. Soc. Amer.* **5**, 494 (1921). (Report on pyrometry.)
(13) F. H. Schofield and D. C. Gall. *J. sci. Instrum.* **1**, 193 (1924). (Optical pyrometer.)
(14) S. Boyer. *Industr. Engng Chem.* **17**, 1252 (1925). (Gallium in quartz thermometers.)
(15) L. Holborn and J. Otto. *Z. InstrumKde*, **46**, 415 (1926). (Mercury in glass thermometers.)
(16) B. Noyes. *J. opt. Soc. Amer.* **13**, 223 (1926). (Clinical thermometers.)
(17) W. F. Higgins. *J. sci. Instrum.* **3**, 361 (1926). (Spirit thermometers.)
(18) W. F. Higgins. *J. R. Soc. Arts*, **74**, 946 (1926). (Cantor Lectures on thermometry.)
(19) F. Henning and C. Tingwald. *Z. Phys.* **48**, 805 (1928). (Flames.)
(20) E. Griffiths and J. H. Awbery. *Proc. roy. Soc.* **123**, 401 (1929). (Flames.)
(21) W. F. Higgins and E. G. Bilham. *Meteorol. Office, Professional Notes*, **51**, (1929). (Mercury thermometers.)
(22) J. A. Beattie, D. D. Jacobus and J. M. Gaines. *Proc. Amer. Acad. Arts Sci.* **66**, 167 (1930). (Construction of resistance thermometers.)
(23) W. T. David, W. Davies and J. Jordan. *Phil. Mag.* **12**, 1043 (1931). (Flame temperatures.)
(24) H. C. Hottel and P. P. Broughton. *Industr. Engng Chem.* **4**, 166 (1932). (A pyrometer for flame temperatures.)
(25) Anon. *Engineer*, **156** (Suppt. Sept. 8), 6 (1933). (Crossed filament pyrometer.)
(26) W. P. White. *Rev. sci. Instrum.* **4**, 142 (1933). (Use of thermocouples.)
(27) A. Shirodkar. *Phil. Mag.* **15**, 426 (1933). (Uses range of α-particles to measure flame temperature.)
(28) J. A. Hall. *J. sci. Instrum.* **10**, 4 (1933). (Construction of resistance thermometers.)

The properties of helium at low temperature.

(29) W. H. Keesom. *Proc. K. Akad. wet. Amst.* **29**, 1136 (1926). (Solid helium.)
(30) W. H. Keesom. *C. R. Acad. Sci.*, Paris, **183**, 26 (1926). (No triple point.)
(31) G. Tamman. *Ann. Phys.*, Lpz., **82**, 240 (1927). (Discusses absence of a triple point.)

(32) W. H. KEESOM and M. WOLFKE. *C. R. Acad. Sci.*, Paris, **185**, 1465 (1927). (Two forms of liquid.)
(33) W. H. KEESOM. *Nature*, Lond., **122**, 847 (1928). (Two forms of liquid.)
(34) W. H. KEESOM and M. WOLFKE. *Proc. K. Akad. wet. Amst.* **31**, 81 (1928). (Two forms of liquid.)
(35) F. SIMON. *Nature*, Lond., April 7, 1934. (Thermodynamic implications.)

On refractories.

(36) *Bur. Stand. J. Res.*, Wash., **6**, 1131 (1931).

Furnaces.

(37) W. FEHRE. *Z. tech. Phys.* **8**, 119 (1927). (Tungsten winding.)
(38) H. S. BOOTH and R. WARD. *Industr. Engng Chem.* **4**, 199 (1932). ("Globar" elements.)
(39) G. BURROWS and F. L. CLARK. *J. sci. Instrum.* **10**, 248 (1933). (Molybdenum winding.)

SECTION 3

(1) E. COHEN and A. L. T. MOESVELD. *Z. phys. Chem.* **100**, 151 (1922). (Adiabatic calorimeter.)
(2) C. V. BOYS. *Engineering*, **115**, 559 (1923). (Gas calorimeter.)
(3) F. H. SCHOFIELD. *J. sci. Instrum.* **1**, 141 (1924). (Calorimeter using mercury as fluid.)
(4) ANON. *Engineering*, **123**, 499 (1927). (Bomb calorimeter.)
(5) F. M. JAEGER and E. ROSENBOHM. *Proc. K. Akad. wet. Amst.* **30**, 905 (1927). (Metal calorimeter.)
(6) E. GRIFFITHS. *Proc. phys. Soc.* **40**, 46 (1928). (Calorimeter for heat developed by fruit.)
(7) E. GRIFFITHS. *Nature*, Lond., **128**, 485 (1931). (Definition of calorie.)
(8) E. J. WORKMAN. *Phys. Rev.* **37**, 1345 (1931). (Continuous flow calorimetry of gases.)
(9) H. R. LANG. *Proc. phys. Soc.* **43**, 572 (1931). (Continuous flow calorimetry of liquids.)
(10) L. ALGERA. *Proc. K. Akad. wet. Amst.* **34**, 906 (1931). (Micro-calorimeter.)
(11) L. G. CARPENTER and T. F. HARLE. *Proc. phys. Soc.* **44**, 389 (1932). (Vacuum calorimeter.)
(12) K. S. COLE and R. H. COLE. *Rev. sci. Instrum.* **3**, 684 (1932). (Dewar flask as calorimeter.)
(13) C. G. HYDE and F. E. MILLS. *Gas Calorimetry* (London, 1932).
(14) A. FERGUSON and J. T. MILLER. *Proc. phys. Soc.* **45**, 194 (1933). (Cooling method.)
(15) W. G. MARLEY. *Proc. phys. Soc.* **45**, 591 (1933). (Cooling method.)
(16) E. D. COONAND and F. DANIELS. *J. phys. Chem.* **37**, 1 (1933). (Micro-calorimeter.)
(17) T. C. SUTTON. *J. sci. Instrum.* **10**, 286 (1933). (Bomb calorimeter.)
(18) H. L. BRONSON, H. M. CHISHOLM and S. M. DOCKERTY. *Canad. J. Res.* **8**, 282 (1933). (Calorimeter of solid copper.)
(19) W. P. WHITE. *The Modern Calorimeter* (New York: Chemical Catalog Co.).

SECTION 4

(1) C. H. LEES. *Philos. Trans.* **192**, 399 (1898). (Bad conductors.)
(2) H. L. CALLENDAR. *Ency. Brit.* (10th ed.), **27**, 193. Article "Conduction of Heat."
(3) W. JAEGER and H. DIESSELHORST. *S.B. preuss. Akad. Wiss.* **37**, 719–26 (1899); *Wiss. Abh. phys.-tech. Reichsanst.*, Berl., **3**, 269 (1900).
(4) C. H. LEES. *Philos. Trans.* **208**, 381 (1908). (Metals.)
(5) E. GRIFFITHS and G. W. C. KAYE. *Proc. roy. Soc.* **104**, 71 (1923). (Poor conductors.)

(6) G. W. C. Kaye and W. F. Higgins. *Proc. roy. Soc.* **117**, 459 (1927). (Liquids.)
(7) A. Eucken and G. Kahn. *Z. phys. Chem.* **134**, 193 (1928). (Low temperature.)
(8) E. Griffiths. *Proc. phys. Soc.* **41**, 151 (1929). (Critical survey.)
(9) C. H. Lees and H. A. Nancarrow. *Phil. Mag.* **14**, 811 (1932). (Rapid method.)
(10) R. W. B. Stevens. *Phil. Mag.* **14**, 897 (1932). (Bad conductors.)
(11) M. Jakob. *Phys. in regelmäss. Ber.* **1**, 123 (1933). (Critical survey.)

Section 5

(1) E. Griffiths. *Methods of Measuring Temperature* (Griffin and Co., 1920).
(2) G. R. Greenslade. *Phys. Rev.* **15**, 150 (1920). (Radiation from metals.)
(3) H. Rubens and G. Michel. *Phys. Z.* **22**, 569 (1921). (Experimental test of Planck's law.)
(4) C. Davisson and J. R. Weeks. *J. opt. Soc. Amer.* **8**, 581 (1924). (Radiation from metals.)
(5) G. N. Lewis. *Proc. nat. Acad. Sci.*, Wash., **13**, 471 (1927). (Generalization of Planck's law.)
(6) W. E. Forsythe. *J. opt. Soc. Amer.* **16**, 307 (1928). (Temperature radiation.)
(7) C. H. Johnson. *Phil. Mag.* **5**, 301 (1928). (Radiation from flames.)
(8) R. Hase. *Phys. Z.* **29**, 904 (1928). (Spectrum of liquid and solid iron.)
(9) C. H. Cartwright. *Rev. sci. Instrum.* **1**, 592 (1930). (The design of vacuum thermopiles.)
(10) R. Hase. *Proc. phys. Soc.* **43**, 212 (1931). (Radiation from metals.)

Section 6

(1) A. H. Davis. *Phil. Mag.* **40**, 692 (1920).
(2) A. H. Davis. *Phil. Mag.* **43**, 329 (1922).
(3) A. H. Davis. *Phil. Mag.* **44**, 920 (1922).
(4) A. H. Davis. *Phil. Mag.* **44**, 940 (1922).
(5) C. W. Rice. *J. Amer. elect. Engrs*, **42**, 653 (1923). (Similitude is applied systematically and Langmuir's film theory adopted.)
(6) C. W. Rice. *Industr. Engng Chem.* **16**, 460 (1924).
(7) A. H. Davis. *Phil. Mag.* **47**, 1057 (1924). (Forced convection.)
(8) S. R. Parsons. *Phys. Rev.* **27**, 788 (1926). (Compares convection at inside and outside of a pipe.)
(9) L. Prandtl. *Phys. Z.* **29**, 487 (1928). (General principles of the boundary layer theory.)
(10) J. Small. *Engineering*, **132**, 569 (1931). (Heat loss as function of angle to wind stream.)
(11) M. ten Bosch. *Phys. Z.* **32**, 31 (1931). (Fluid in tubes.)
(12) W. S. Kimball and W. J. King. *Phil. Mag.* **13**, 888 (1932).
(13) W. S. Kimball and L. D. Childs. *Phil. Mag.* **14**, 337 (1932). (Semi-theoretical treatment.)
(14) W. S. Kimball and W. J. King. *Phil. Mag.* **14**, 570 (1932). (Semi-theoretical.)
(15) N. Fishenden and O. Saunders. *The Calculation of Heat Transmission* (H.M. Stationery Office, 1932).
(16) W. H. McAdams. *Heat Transmission* (McGraw Hill Book Co., 1933).
(17) E. Griffiths and J. H. Awbery. *Proc. Instn Mech. Engrs*, **125**, 319 (1933). (Convection at surface of pipes.)
(18) W. Findeisen. *Beitr. Geophys.* **39**, 356 (1933). (Relation to mechanical force, on the boundary layer theory.)
(19) N. A. V. Piercy and H. F. Winny. *Phil. Mag.* **16**, 390 (1933). (Theory for flat plate.)
(20) J. T. McIntyre. *J. R. tech. Coll. Glasg.* **3**, 86 (1933). (Fluids in tubes.)
(21) E. Hofmann. *Phys. Z.* **34**, 208 (1933). (Fluid in a tube.)

SECTION 7

(1) P. DEBYE and E. HÜCKEL. *Phys. Z.* **24**, 305 (1923).
(2) E. HÜCKEL. *Phys. Z.* **26**, 93 (1925).
(3) A. FERGUSON and I. VOGEL. *Phil. Mag.* **1**, 971 (1925). (Test of Debye-Hückel theory.)
(4) A. FERGUSON and I. VOGEL. *Phil. Mag.* **4**, 233 (1927). (Test of Debye-Hückel theory.)
(5) A. FERGUSON and I. VOGEL. *Phil. Mag.* **4**, 300 (1927). (Test of Debye-Hückel theory.)
(6) W. F. GIAUQUE and H. L. JOHNSTON. *J. Amer. chem. Soc.* **50**, 3220 (1928).
(7) T. DE DONDER. *C. R. Acad. Sci.*, Paris, **186**, 1599 (1928). (Relativity.)
(8) T. W. RICHARDS and M. DALE. *J. Amer. chem. Soc.* **50**, 794 (1929). (Test of Debye-Hückel theory.)
(9) G. N. LEWIS. *J. Amer. chem. Soc.* **53**, 2578 (1931). (Discussion of the second law.)
(10) F. G. DONNAN and E. A. GUGGENHEIM. *Z. phys. Chem.* A, **162**, 346 (1932). (Membrane equilibria.)
(11) T. S. WHEELER. *Phil. Mag.* **13**, 604 (1932). (Strong electrolytes.)
(12) C. W. DAVIES. *The Conductivity of Solutions* (London, 1933).
(13) R. C. TOLMAN and H. P. ROBERTSON. *Phys. Rev.* **43**, 564 (1933). (Relativity.)
(14) R. C. TOLMAN. *Science*, **77**, 291 (1933).
(15) R. C. TOLMAN. *Science*, **77**, 313 (1933).
(16) R. C. TOLMAN. *Relativity, Thermodynamics and Cosmology* (Oxford Press, 1934).

SECTION 8

(1) W. H. KEESOM. See references in section 2.
(2) M. JAKOB. *Z. Ver. dtsch. Ing.* **65**, 568 (1921). (Properties of steam.)
(3) *Bur. Stand. Circular*, No. 142 (1923). (Properties of ammonia.)
(4) O. KNOBLAUCH. *Z. tech. Phys.* **6**, 638 (1925). (Properties of steam.)
(5) K. BENNEWITZ and E. SPLITTGERBER. *Z. phys. Chem.* **124**, 49 (1926). (Specific heat at critical point.)
(6) E. P. BARTLETT. *J. Amer. chem. Soc.* **49**, 65 (1927). (Properties of gases containing water.)
(7) K. BENNEWITZ and H. WENDROTH. *Z. phys. Chem.* **125**, 111 (1927). (Critical state.)
(8) E. SCHRÖER. *Z. phys. Chem.* **129**, 79 (1927). (Critical state.)
(9) H. CALLENDAR. *J. R. Soc. Arts* (1927). (The Howard Lectures.)
(10) J. A. BEATTIE and O. C. BRIDGMAN. *J. Amer. chem. Soc.* **50**, 3133 (1928). (Equation of state.)
(11) H. CALLENDAR. *Engineering*, **126**, 530 *et seq.* (1928).
(12) H. CALLENDAR. *Proc. roy. Soc.* **120**, 460 (1928). (Properties of steam.)
(13) R. PLANCK and E. KUPRIANOFF. *Z. tech. Phys.* **10**, 93 (1929). (Properties of carbon dioxide.)
(14) M. WOLFKE and J. MAZUR. *Z. phys. Chem.* **74**, 110 (1932). (Two forms of liquid.)

SECTION 9

(*a*) *Vapour pressure and its measurement.*

(1) L. HOLBORN and A. BAUMANN. *Ann. Phys.*, Lpz., **31**, 945 (1910). (Dead-weight gauge.)
(2) F. G. KEYES and R. B. BROWNLEE. *J. Amer. chem. Soc.* **40**, 25 (1918). (Dead-weight gauge.)
(3) I. BENCOWITZ and H. T. HOTCHKISS. *J. phys. Chem.* **30**, 643 (1926). (Vapour pressure of solutions.)
(4) W. EDWARDSON and A. EGERTON. *Proc. roy. Soc.* **113**, 520 (1927).
(5) H. MAYER. *Z. Phys.* **67**, 240 (1930). (A method for pressure as low as 10^{-7} mm. mercury.)
(6) *Bur. Stand. J. Res.*, Wash., **6**, 1061 (1931). (Dead-weight gauge.)

(*b*) *Evaporation and condensation.*

(7) A. Ferguson and N. Thomas. *Phil. Mag.* **34**, 308 (1917). (Data.)
(8) W. Gallenkamp. *Monthly Weather Review*, **47**, 857 (1919).
(9) F. Carrier. *Industr. Engng Chem.* **13**, 432 (1921). (Data.)
(10) T. B. Hine. *Phys. Rev.* **24**, 79 (1924). (Data for many liquids.)
(11) I. S. Bowen. *Phys. Rev.* **27**, 779 (1926). (Evaporation from lakes, etc.)
(12) J. S. Owens. *Proc. roy. Soc.* **110**, 378 (1926). (Condensation on hygroscopic nuclei.)
(13) E. Schmidt. *Gesundheitsing.* **52**, 526 (1929). (Similitude.)
(14) W. Nusselt. *Z. angew. Math. Mech.* **10**, 105 (1930). (Similitude.)
(15) T. Alty. *Proc. roy. Soc.* **131**, 554 (1931). (Mechanism.)
(16) T. Alty and F. H. Nicholl. *Canad. J. Res.* **4**, 547 (1931). (Mechanism.)
(17) A. Svenssen. *Ark. Mat. Astr. Fys.*, **22** A (1932). (Data.)
(18) M. Jakob. *Z. Ver. dtsch. Ing.* **76**, 1161 (1932).
(19) S. K. Banerji and H. M. Wodia. *Mem. Indian met. Dep.* **25**, 261 (1932). (Mechanism.)
(20) T. Alty. *Phil. Mag.* **15**, 82 (1933). (Maximum rate of evaporation of water.)
(21) E. Preston. *Trans. Faraday Soc.* **29**, 1188 (1933). (Theory.)
(22) K. Schroeber. *Z. tech. Phys.* **14**, 81 (1933). (Survey.)
(23) W. Findeisen. *Beitr. Geophys.* **39**, 356 (1933). (Relation to mechanical force.)
(24) *Rep. nat. phys. Lab.* **33** (1933).

Section 10

(1) J. Cortus. *Z. Ver. dtsch. Ing.* **65**, 767 (1921). (Thermocouple psychrometer.)
(2) E. Griffiths. *The Measurement of Humidity in Closed Spaces* (H.M. Stationery Office).
(3) D. C. Rose. *Canad. J. Res.* **5**, 156 (1931). (Resistance thermometer psychrometer.)
(4) L. Kettenacker. *Z. InstrumKde*, **52**, 319 (1932). (Thermocouple psychrometer with very fine wires.)
(5) P. le Rolland and T. T. Lou. *C. R. Acad. Sci.*, Paris, **194**, 258 (1932). (Method of observing dew point.)
(6) J. H. Awbery and E. Griffiths. *Proc. phys. Soc.* **44**, 132 (1932). (Psychrometry at high temperatures.)
(7) M. Robitzsch. *Z. InstrumKde*, **52**, 80 (1932). (Theory of psychrometry.)
(8) N. E. Edlefsen. *Rev. sci. Instrum.* **4**, 345 (1933). (Electrical resistance of glass-wool hygrometer.)
(9) J. H. Arnold. *Physics*, **4**, 255 (1933). (Theory of psychrometer.)
(10) J. H. Arnold. *Physics*, **4**, 334 (1933). (Theory of psychrometer.)

Section 11

(1) M. Knudsen. *Ann. Phys.*, Lpz., **34**, 593 (1911). (Conduction in gases.)
(1*a*) M. Knudsen. *Ann. Phys.*, Lpz., **28**, 104 (1909). (Cosine law.)
(2) I. Langmuir. *Phys. Rev.* **8**, 149 (1916). (Evaporation and condensation.)
(3) A. Eucken. *Verh. dtsch. phys. Ges.* **18**, 18 (1916). (Specific heat of hydrogen.)
(4) R. C. Tolman. *J. Amer. chem. Soc.* **43**, 1592 (1921). (Entropy of electron gas, assuming Maxwell statistics.)
(4*a*) Ward. *Phil. Mag.* **32**, 365 (1916).
(5) J. Choriton and N. Semenoff. *Z. Phys.* **25**, 287 (1924). (Critical temperature for condensation.)
(6) E. Fermi. *Z. Phys.* **36**, 902 (1926). (Extension of Pauli's principle to gases.)
(7) M. Knudsen. *Ann. Phys.*, Lpz., **83**, 797 (1927). (Thermo-molecular pressures.)
(8) J. L. Costa, H. D. Smyth and K. T. Compton. *Phys. Rev.* **30**, 349 (1927). (Test of Maxwell's distribution law.)
(9) J. A. Eldridge. *Phys. Rev.* **30**, 931 (1927). (Test of Maxwell's distribution law.)
(10) M. Knudsen. *Ann. Phys.*, Lpz., **6**, 129 (1930). (Accommodation coefficients.)
(11) W. R. van Wijk. *Proc. K. Akad. wet. Amst.* **34**, 1201 (1931). (Accommodation coefficients.)

(12) C. H. CARTWRIGHT. *Physics*, **1**, 211 (1931). (Limitations on radiometers due to Brownian movement.)
(13) K. B. BLODGETT and I. LANGMUIR. *Phys. Rev.* **40**, 78 (1932). (Accommodation coefficients.)
(14) J. M. JACKSON and N. F. MOTT. *Proc. roy. Soc.* **137**, 703 (1932). (Accommodation coefficients discussed by wave mechanics.)

SECTION 12

(1) A. EUCKEN and O. NEUMANN. *Z. phys. Chem.* **111**, 431 (1924). (Lorentz function—double mechanism.)
(2) A. EUCKEN and K. DITTRICH. *Z. phys. Chem.* **125**, 211 (1927). (Lorentz function.)
(3) A. SOMMERFELD. *Z. Phys.* **47**, 1 (1928). (Theory of metals.)
(4) A. EUCKEN and G. KAHN. *Z. phys. Chem.* **134**, 193 (1928). (Low temperature results for salts.)
(5) C. C. BIDWELL. *Phys. Rev.* **32**, 311 (1928). (Double mechanism postulated.)
(6) R. W. B. STEVENS. *Phil. Mag.* **14**, 897 (1932). (Pyrex glass.)

SECTION 13

(1) M. BORN and TH. VON KÁRMÁN. *Phys. Z.* **14**, 65 (1913). (Crystal lattice.)
(2) E. SCHRÖDINGER. *Phys. Z.* **20**, 497 (1919). (Born's theory.)
(3) M. BORN and E. BRODY. *Z. Phys.* **11**, 327 (1922). (Crystal lattice.)
(4) J. A. BEATTIE. *J. Math. Phys.* **6**, 1 (1926). (Tables of Debye function.)
(5) K. CLUSIUS and P. HARTECK. *Z. phys. Chem.* **134**, 243 (1928). (Low temperature.)
(6) Y. ENDO. *Rep. aero. Res. Inst. Tokyo*, **71**, 71 (1931). (Lattice solid.)
(7) W. H. KEESOM and J. VAN DEN ENDE. *Proc. K. Akad. wet. Amst.* **35**, 143 (1932). (Discontinuity at transition point.)
(8) W. H. KEESOM and J. A. KOK. *Proc. K. Akad. wet. Amst.* **35**, 301 (1932). (Silver.)
(9) H. L. BRONSON, H. M. CHISHOLM and S. M. DOCKERTY. *Canad. J. Res.* **8**, 282 (1933). ($Cp = A + BT - C/T^3$.)

SECTION 14

(1) O. KNOBLAUCH and E. RAISCH. *Z. Ver. dtsch. Ing.* **66**, 418 (1922). (Superheated steam.)
(2) E. C. KEMBLE and J. H. VAN VLECK. *Phys. Rev.* **21**, 653 (1923). (Allowance for expansion under centrifugal force.)
(3) J. R. PARTINGTON and W. G. SHILLING. *The Specific Heats of Gases* (London, 1924).
(4) H. C. HICKS and A. C. G. MITCHELL. *J. Amer. chem. Soc.* **48**, 1520 (1926). (Old quantum theory.)
(5) W. H. MCCREA. *Proc. Camb. phil. Soc.* **23**, 942 (1927).
(6) E. HUTCHISSON. *J. Amer. chem. Soc.* **50**, 1895 (1928). (Weights correctly assigned.)
(7) W. F. GIAUQUE and H. L. JOHNSTON. *J. Amer. chem. Soc.* **50**, 3221 (1928). (Properties of ortho- and para-hydrogen.)
(8) W. F. GIAUQUE. *J. Amer. chem. Soc.* **52**, 4808 (1930). (Para- and ortho-hydrogen—a useful summary.)
(9) W. F. GIAUQUE. *J. Amer. chem. Soc.* **53**, 507 (1931). (Effect of nuclear spin.)
(10) H. O. KNESER. *Ann. Phys.*, Lpz., **11**, 761, 777 (1931) and 1015 (1932).
(11) A. R. GORDON and C. BARNES. *J. phys. Chem.* **36**, 2292 (1932).
(12) R. M. BADGER and SHO-CHOW-WOO. *J. Amer. chem. Soc.* **54**, 3523 (1932). (Polyatomic linear molecules.)
(13) A. R. GORDON and C. BARNES. *J. chem. Phys.* **1**, 297 (1933). (Valuable tables and hints on technique.)
(14) H. L. JOHNSTON and A. T. CHAPMAN. *J. Amer. chem. Soc.* **55**, 153 (1933). (Triatomic molecules.)
(15) L. S. KASSEL. *J. Amer. chem. Soc.* **55**, 1351 (1933). (Polyatomic molecules.)
(16) G. W. C. KAYE and G. G. SHERRATT. *Proc. roy. Soc.* **141**, 123 (1933). (Sound-velocity method.)

ELECTRICAL PHENOMENA AT EXTREMELY LOW TEMPERATURES*

By Prof. J. C. McLennan, D.Sc., LL.D., F.R.S.

§ 1. INTRODUCTION

With the liquefaction of the gas helium, and its final solidification by Prof. Keesom[1] on the 26th June, 1926, the last page might be said to have been written to a chapter of brilliant achievement in science; a chapter whose opening paragraphs record the experiments and discoveries of Michael Faraday at the Royal Institution in the early nineteenth century[2].

It was Lavoisier, the great chemist, who suggested to his doubtless sceptical contemporaries that if the earth were cooled sufficiently, even the gaseous air above might take a liquid form, but many years were to pass before the liquefaction of gases was demonstrated in the laboratory. Sir Humphry Davy and his young assistant Faraday found that the gas chlorine, formed by heating its hydrate in a sealed tube, could be condensed to a yellow liquid. Ammonia, sulphuretted hydrogen, and hydrochloric acid gas, quickly yielded to their experimental skill, and by 1823 Faraday had installed at the Royal Institution what might truly be called a liquefaction plant.

Progress in both the technique and the range of application of such experiments was rapid. The little cylinders we use to recharge the soda-water siphons, and the "dry ice" or "snow" used extensively in hospitals and for commercial refrigeration purposes, have made us familiar to-day with carbon dioxide, normally a gas, in its liquid and its solid forms; but they were introduced to the world in 1834 by Thilorier. Faraday was quick to turn to advantage the new discoveries; by mixing carbon-dioxide snow with ether he obtained a freezing mixture, immersed in which his test tubes might be cooled to a temperature of $-110°$ C., or, on the commonly used Fahrenheit scale, 166 degrees below zero; a temperature far below any occurring naturally on the earth. A field of research full of new and strange possibilities opened up before the physicists and chemists of the day. Would even the gases of the atmosphere, oxygen and nitrogen, would hydrogen and the other gases that had so far resisted all attempts to liquefy them, yield to the new low temperatures now reached by science, or were they really "permanent" gases as so many thought? Thomas Andrews, the professor of chemistry at Belfast, tried to effect the liquefaction, and, like others before him, failed, yet made of his very failure a success. In one of the long-distinguished Bakerian lectures before the Royal Society he explained his experiments with carbon dioxide (his original apparatus is to-day in the South Kensington Museum), which showed that in the

* The Twenty-fifth Kelvin Lecture delivered before the Institution of Electrical Engineers, April 26, 1934.

liquefaction of a gas not only the temperature but also the pressure to which it was subjected played an essential part. Here was the key to the locked door; unless the pressure exceeds a certain "critical pressure," and the temperature is below a certain "critical temperature," peculiar to the gas, we can never hope to liquefy it; cold allied to compression succeeds where the one alone fails.

The story then is one of triumph after triumph, as one gas after another fell before the advance of scientific achievement. Oxygen was liquefied by Olszewski in 1883, and two years later nitrogen and carbon monoxide had proved to be non-permanent gases after all, so that the "liquid air" of Lavoisier's dream had become an accomplished fact. The list of lowest temperatures as they were reached reads like the records of athletic achievement; in both the next step is always harder to achieve than the one before. A new principle was applied to produce the cooling, a principle that is indeed the basis of our modern methods of refrigeration. Anyone who has used a bicycle pump must have noticed that compression of a gas produces heat and raises its temperature; that the converse is equally true is perhaps not so well known. If a gas under pressure below its inversion temperature is allowed to expand freely through a fine nozzle or if it is allowed to do mechanical work while expanding, it will cool, and may reach a temperature at which it liquefies. By an ingenious arrangement of the apparatus, the gas cooled by such a process is made to lower the temperature of the oncoming gas, which on expansion reaches a still lower temperature, until at last the liquid is produced. By such means as these, hydrogen was at length liquefied by Olszewski in 1892, and there remained of all known substances only the gas helium as yet uncondensed.

Hydrogen is a liquid at $-253°$ C., or only 20 degrees above what we call the absolute zero of temperature, but these remaining 20 degrees present difficulties of attainment vastly greater than 20 degrees of normal temperature. Practically every factor that conspires to prevent further cooling has here increased in importance. It is only by the use of the vacuum flasks invented by Dewar that the access of heat from the surroundings is prevented; indeed, without his invention, all progress in liquefaction must have reached an early limit. Any foreign gas existing in the helium as an impurity must be removed; its presence would mean that the apparatus would soon be choked with solid substances, for at $-259°$ C. even hydrogen has solidified. It was at Leiden in 1908 in the laboratories of Prof. Kamerlingh Onnes that at length the goal was reached. At a temperature of $-269°$ C., or about 4·2 degrees above absolute zero, the first drop of liquid helium fell into the receiving flask; very similar in appearance to a common drop of water, but the climax of a hundred years' achievement. Few in London have had an opportunity of seeing this rare and hardly-won liquid except those who were present at my lectures in June, 1932, and January, 1933, at the Royal Institution.

It is a long way from Michael Faraday's comparatively simple experiments to the operation of a modern liquid-helium plant, but just as to produce liquid helium we first make liquid air, then by its use cool the apparatus to make liquid hydrogen, and with this make possible the liquefaction of helium; just so has each achievement been built upon the work preceding it.

By reducing the pressure over liquid helium, still lower temperatures can be reached; and it was in this way that Keesom succeeded about a year ago in cooling liquid helium to an estimated temperature of about 0·71 degree above the absolute zero. In 1926 Keesom[1] also succeeded in demonstrating at Leiden the solidification of helium by applying high pressure to helium in the liquid state. Thus man has conquered, and every known substance has yielded to his will; at such temperatures as are now attainable, all the world can be solidified, and gases and liquids will exist no more.

Liquid helium is at present available for research in such centres as Leiden; Toronto; Berlin; Breslau; Washington; Berkeley, California; Oxford; and Cambridge.

§ 2. MEASUREMENT OF TEMPERATURES

It is possible by means of observations and an application of the gas laws to determine equations of state for helium that enable one to establish a temperature scale applicable over a range extending down to about 1·5° K. Below this temperature we are largely dependent on extrapolation. The vapour pressure of liquid helium enables us to reach the temperature 0·71° K., and the magnetic susceptibilities of certain salts can be used to tell us when such temperatures as 0·05° K. are reached.

Table 1. Vapour pressure of liquid helium[4]

Vapour pressure (cm. of Hg)	Temperature (°K.)	Vapour pressure (cm. of Hg)	Temperature (°K.)
132·9	4·899	0·500	1·538
76·00	4·219	0·250	1·393
50·00	3·795	0·1000	1·237
25·00	3·218	0·0500	1·139
10·00	2·636	0·0250	1·055
5·00	2·298	0·0100	0·960
2·50	2·013	0·0050	0·899
1·000	1·714	0·0025	0·844

For the study of properties of matter between the temperatures of 4·2° K. and 0·05° K. a liquid-helium cryostat is uniquely suitable. The temperature of the liquid helium can be modified by varying its equilibrium vapour pressure, and readings on the vapour pressure can easily be made with an accuracy of at least $\frac{1}{100}$ degree. That this is so can be seen from the numbers recorded in table 1.

§ 3. LIQUID HELIUM

Liquid helium is a light, colourless, mobile liquid. Its density varies with its temperature and with the pressure to which it is subjected. Under a pressure of 1 atmosphere and at a temperature of 4·2° K., it is 0·1252 or about one-eighth the density of water. A series of densities measured for liquid helium by Prof. and Miss Keesom[5] is given in table 2. For each pressure to which it is subjected, liquid helium reaches a maximum density at a definite and characteristic temperature,

known as the λ point. For a pressure of 1 atmosphere this temperature is 2·178° K., and for a pressure of 25 atmospheres it is 1·835° K. The corresponding maximum densities are respectively 0·1473 and 0·1757. The higher the pressure the lower is the λ point, and the greater the critical density.

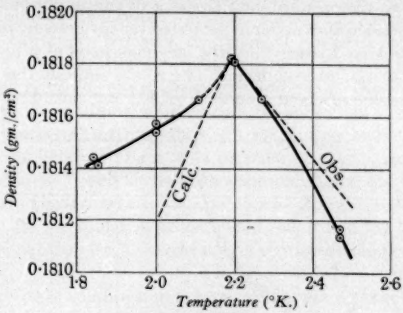


Figure 1.

A curve of densities taken from a paper by Keesom[6] is shown in figure 1. This curve consists of two branches which meet at the λ point under a finite angle. Above this point the liquid is known as helium I, and below as helium II[7]. Immediately above the λ point the coefficient of expansion for temperature is positive and is given by $\alpha_1 = 0{\cdot}0222$ deg.$^{-1}$, and just below it is negative and given by $\alpha_{11} = -0{\cdot}0426$ deg.$^{-1}$. Helium I and helium II are both highly compressible.

Table 2. Density of liquid helium (gm./cm.³)

Temp. (°K.)	$p=1$	$p=5$	$p=10$	$p=15$	$p=20$	$p=25$	$p=30$	$p=35$ atm.
1·25	0·1462	0·1522	0·1584	0·1636	0·1681	0·1722	—	—
1·50	0·1463	0·1524	0·1585	0·1638	0·1686	0·1728	—	—
1·75	0·1465	0·1527	0·1591	0·1646	0·1696	0·1743	—	—
1·80	0·1465	0·1528	0·1592	0·1648	0·1699	0·1750	0·1796	—
1·90	0·1466	0·1531	0·1597	0·1655	0·1711	0·1757	0·1795	—
2·00	0·1468	0·1534	0·1604	0·1666	0·1713	0·1756	0·1794	0·1829
2·10	0·1470	0·1540	0·1609	0·1665	0·1712	0·1755	0·1793	0·1828
2·20	0·1472	0·1540	0·1608	0·1663	0·1710	0·1753	0·1790	0·1826
2·25	0·1471	0·1539	0·1606	0·1662	0·1709	0·1752	0·1789	0·1824
2·50	0·1460	0·1530	0·1600	0·1655	0·1703	0·1745	0·1783	0·1818
3·00	0·1424	0·1505	0·1579	0·1637	0·1686	0·1730	0·1768	0·1803
3·50	0·1370	0·1469	0·1551	0·1615	0·1666	0·1712	0·1751	0·1786
4·00	0·1292	0·1421	0·1518	0·1587	0·1644	0·1691	0·1732	0·1768
4·20	0·1252	0·1399	0·1502	0·1575	0·1634	0·1682	0·1724	0·1761
	λ curve						Solidification curve	
ρ	0·1473	0·1541	0·1610	0·1666	0·1714	0·1757	0·1796	0·1831
T	2·178	2·125	2·058	1·990	1·915	1·835	1·757	1·920

A provisional determination made by Keesom and Clusius[8] of the compressibility of liquid helium has given for χ at 2·71° K. a value of $7{\cdot}18 \times 10^{-3}$ cm.² per kg. for $p = 8{\cdot}87$ kg. per cm.² This result is interesting because it is the highest value obtained as yet for the compressibility of a liquid. The dielectric constant of liquid helium varies with temperature, and shows a discontinuity at the λ point. The specific heat of liquid helium under its saturated vapour pressure has been measured also by Prof. and Miss Keesom[9], in the neighbourhood of 2·19° K., the λ point, with increments of 0·01 to 0·02 degree. The results indicate that the specific heat has a maximum value of about 3·00 at 2·19° K., and falls down to 1·1 certainly

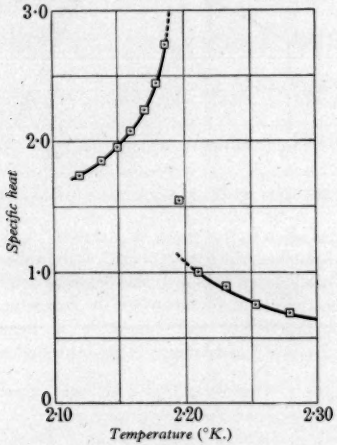


Figure 2.

within 0·02 degree and probably within 0·002 degree. This discontinuity in the value of the specific heat is shown clearly by the curve in figure 2, where the ordinates are expressed in calories per degree K. per gramme.

From these and other characteristics of liquid helium it is clear that in helium I and helium II we have two phases, each possessing clearly defined characteristic properties that will bear further investigation. The heat conductivity, the viscosity, the refractivity, as well as the structure and composition, for example, of liquid helium, might be studied with advantage. It is probable that phases similar to those of helium I and helium II will be found to characterize liquefied neon, argon, krypton, and xenon, and as these are more easily obtainable from a physical point of view than liquid helium, they might well be given consideration.

§ 4. SOLIDIFIED HELIUM

For what information we possess regarding the solidification of helium, we are indebted to Keesom and Simon, with their associates. As already mentioned, helium was first solidified by Keesom[9] in 1926. Kamerlingh Onnes had attempted to solidify liquid helium by reducing, by means of powerful pumps, the vapour pressure to which it was subjected. He succeeded in reaching a temperature of 0·8° K. but without obtaining any evidence of solidification. Keesom, reversing the process, succeeded in obtaining solidification by subjecting liquid helium to a series of high pressures. At a temperature of 1·10° K. a pressure of 26 kg. per cm^2 sufficed, but at a temperature of 4·2° K. a pressure of 145 kg. per cm^2 was found to be necessary. Simon[10] and his associates showed that at 800 atmospheres the melting point of solid helium was 12° K., and that at 5600 atmospheres the melting point was 40° K. By extrapolation he estimated that to solidify helium at a temperature of 100° K. (or −173·2° C.) a pressure of approximately 22,000 atmospheres would be required.

The heat of fusion of solid helium is 6·75 calories per gramme-atom at 4·0° K., and 5·1 calories per gramme-atom at 3·4° K. The density in equilibrium with the liquid phase at 4·0° K. is 0·23, and at 3·6° K. it is 0·22. The compressibility, according to Kaischew and Simon[11], is approximately $1{\cdot}5 \times 10^{-3}$ per atmosphere at 3·7° K. and 115 atmospheres.

§ 5. NEW METHODS

(i) *Liquefaction of helium.* For the liquefaction of helium, the method adopted until recently involved the use of the Joule-Thomson effect. When it becomes desirable, however, to produce liquid helium in small quantities and at a low cost, apparatus of the type devised by Simon and Mendelssohn may be used. It is this type of equipment that is in use at Oxford[12]. The method, which is a discontinuous process, involves the compression of helium to about 150 atmospheres in the annular space *B* between the two concentric cylinders shown in figure 3. The experimental material to be studied is placed in the metal container *A*, which is immersed in liquid hydrogen boiling in *D* under reduced pressure.

In the process the container *V* is at first filled with gaseous hydrogen or helium in order to provide a conducting medium to bring the inner cylinder *B* containing the compressed helium to 15° K., the temperature of the liquid hydrogen. For additional cooling, liquid hydrogen boiling under reduced pressure is held in the annular space *C*. When temperature equilibrium is attained the container *V* is evacuated with a powerful diffusion pump, and the helium compressed in *B* is allowed to expand. In this process, about one-half of the helium liquefies and serves to cool the two cylinders *B* and the material in them to the temperature of liquid helium.

By another process, a separate container within the chamber *B*, but not shown in the figure, is filled with charcoal, zeolithe, chabasite, or silica-gel, and allowed to

absorb helium while temperature equilibrium is being established, the chambers A and B being again insulated thermally as before by the evacuation of the space V. The helium absorbed by the charcoal is quickly pumped off and cooling follows. This lowering of the temperature liquefies a part of the helium in the chamber B, and the material in A is consequently cooled to liquid-helium temperatures.

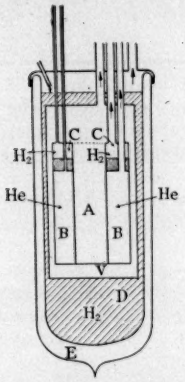


Figure 3. Apparatus of Simon and Mendelssohn.

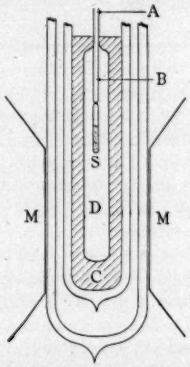


Figure 4. Apparatus used by de Haas to produce the lowest temperatures by demagnetization of salts.

Methods for producing a lowering in temperature under very special conditions have recently been brought forward that involve the adiabatic expansion of solid helium and of liquid helium I, i.e. liquid helium above the λ point. In order to bring about a cooling effect with liquid helium II, it would appear that adiabatic compression rather than expansion would be required.

(ii) *Cooling by demagnetization.* What promises to be a procedure of great importance in research at the lowest temperatures is based on the use of the cooling effect obtainable under certain conditions by adiabatic demagnetization. Debye[13] and also Giauque[14] in 1926 showed that it would be possible to get an effect of a measurable order with paramagnetic salts if the initial temperature were sufficiently low. Giauque and Clarke[15] even suggested that the temperature obtainable in this way might be extremely low.

Recently the method was tested by de Haas, Wiersma, and Kramers[16], and found to be highly efficient. A sketch of the apparatus used is shown in figure 4. The sample of paramagnetic salt S to be magnetized and demagnetized was held in a small glass tube B, which formed part of a small Dewar flask that was suspended from a rod and attached to one arm of a balance. The whole apparatus was placed with the salt suspended between the poles of an electromagnet and occupying the region where $H\partial H/\partial x$ had a maximum value. The suspended Dewar flask was

surrounded, as shown in the diagram, with liquid helium boiling at a temperature of 1·26° K. A small amount of helium gas having a pressure of 10^{-7} mm. of mercury was left in the space *D* to give thermal contact with the liquid helium in the outer vessel *C*. In the initial experiments a field as high as 30,000 gauss was maintained by the magnet until the weight of the system suspended from *A* indicated that an equilibrium state had been reached. One could be certain then that under these conditions the specimen *S* was at the temperature 1·26° K. and that it was exhibiting the magnetic susceptibility corresponding to that temperature. When these conditions obtained and the magnetic field was suddenly reduced to a fixed low value, of 1000 or 500 gauss, for example, the adiabatic demagnetization of the salt *S* ensued, and the lowering of its temperature followed. The salt itself indicated by its susceptibility the temperature reached as a result of the demagnetization, for the curve showing the relation between the pull of the field and the temperature of the

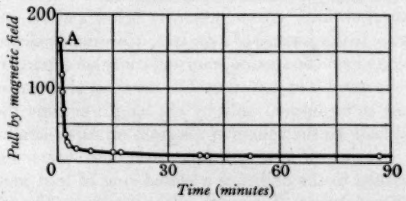


Figure 5. Example of temperature measurement by force exerted on a paramagnetic salt in a magnetic field. Curve shows increase of temperature of salt after demagnetization from $T=A$ to $T=2{\cdot}16°$ K.

specimen had been previously obtained for a range of temperature from 7·2 to 1·3° K. and extrapolated linearly to lower temperatures. The value of the lowest temperature reached as a result of the demagnetization was given by the value of the pull of the field (1000 or 500 gauss) immediately following demagnetization. Conduction of heat to the specimen from the surrounding liquid helium caused the temperature gradually to rise and the pull to lessen owing to the susceptibility falling off. A curve showing this decrease in pull is given in figure 5. This curve calls attention to one defect in the method. It takes, it will be seen, a long time for equilibrium to be reached. This makes the taking of measurements rather tedious, for it operates in both the initial and final stages of the experiment.

The salts used so far have been cerium fluoride [CeF_3], dysprosium ethyl sulphate [$Ds\,(C_2H_5SO_4)_3$ 9 aq.], cerium ethyl sulphate [$Ce\,(C_2H_5SO_4)_3$ 9 aq.], and potassium-chrome alum. With cerium fluoride a temperature of 0·13° K. was reached, with dysprosium ethyl sulphate 0·12° K., with cerium ethyl sulphate 0·085° K., and with potassium-chrome alum 0·03° K., the last-mentioned being the lowest temperature ever attained and measured. It may be added that with the use of manganese ammonium sulphate Kurti and Simon[17] reached a temperature of 0·1° K.

§ 6. LOW-TEMPERATURE ELECTRICAL CONDUCTIVITY OF METALS

General. It is of interest to find that the solutions of many problems connected with the conduction of electricity through metals are likely to be reached through low-temperature research.

The classical experiments of Dewar and Fleming[18] made in 1893 on the electrical conductivity of metals cooled to very low temperatures by means of liquefied gases, including air and hydrogen, come to mind. These yielded results suggesting that the electrical resistances of all pure metals would vanish at the absolute zero of temperature.

This suggestion, however, proved to be wrong, for in 1911, Kamerlingh Onnes at Leiden, while carrying out researches at low temperatures with the aid of liquefied helium, discovered that mercury, when cooled down and solidified with liquid helium, suddenly and abruptly at about 4·2° K. became what is now designated a super-conductor of electricity. At temperatures below 4·2° K., mercury offers no measurable resistance to the passage of a current. Currents of electricity started in a ring of a metal in the super-conducting state will continue apparently undiminished in intensity while the metal is in that state. The duration of these persistent induced ring currents seems to be limited only by the length of time the cooling agent, liquid helium, will last. In the course of a lecture on super-conduction in metals, delivered on the evening of the 3rd June, over two years ago, at the Royal Institution, I exhibited to the audience a closed ring of lead immersed in liquid helium and carrying a current of more than 200 amperes. The current had been started in the super-conducting lead ring some 6 hours earlier in the afternoon by Prof. Keesom in Leiden, and had persisted undiminished in intensity while being transported in liquid helium by aeroplane from Leiden to London by Colonel the Master of Sempill (now Lord Sempill).

Super-conducting metals. Other metals in addition to mercury and lead that exhibit the super-conducting property if made sufficiently cold are tin, indium, vanadium, aluminium, zinc, gallium, thallium, tantalum, titanium, thorium, and niobium. The transition temperature for the passing of a metal from the ordinary conducting to the super-conducting state is not a constant but varies with the metal. For mercury it is 4·22° K., for lead 7·2° K., tin 3·7° K., tantalum 4·4° K., thallium 2·37° K., indium 3·37° K., gallium 1·05° K., thorium 1·5° K., titanium 1·75° K., niobium 9·2° K., vanadium 4·3° K., aluminium 1·14° K., zinc 0·79° K., and cadmium 0·6° K.[17]

Vanadium is especially interesting because next to aluminium and titanium it is the super-conductor with the smallest atomic weight, and unlike the others it has a convenient transition temperature.

Special investigation has shown that silver, magnesium, and tungsten, are still not super-conducting at 0·74° K.; and bismuth, iron, and nickel, not at 0·75° K. Lead remains super-conducting still at 0·75° K.

Alloys and chemical compounds. Some alloys and chemical compounds of the metals also exhibit the super-conducting property. Copper sulphide, for example,

does so, though neither of the constituent elements is a super-conductor. The nitrides and carbides, borides and silicides of several of the metals—such, for example, as those of molybdenum, tungsten, tantalum, zirconium, and niobium—are also super-conducting at sufficiently low temperatures.

The addition of metals of the bismuth group to super-conducting metals has been found, speaking generally, to raise their transition temperature. Bismuth added to lead raises the transition temperature from 7·2° to 8·8° K.; carbon raises that of niobium from 9·2° to 10·1° K. Gold alloyed with bismuth becomes super-conducting at 1·94° K., whereas neither constituent alone becomes super-conducting even at the lowest temperatures obtainable.

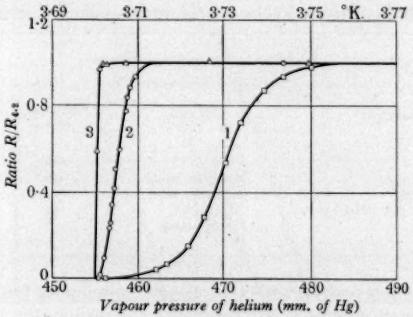


Figure 6.

With pure metals the transition from the ordinary conducting to the super-conducting state generally occurs within a tenth or at most a few hundredths of a degree. With impure metals, alloys, or chemical compounds, the transition is not generally so rapid. The temperature interval for the transition stage has been found by de Haas[19] in the case of tin to be a minimum when the crystalline state of the metal is of the simplest type. This is illustrated by the curves in figure 6, where curve 1 is the transition curve for a wire of poly-crystalline tin, curve 2 is for a tin wire made up of several large crystals, and curve 3 is for a wire made of a single crystal of tin. In the transition stages, for most of the metals the variation of resistance can be readily followed by observing the vapour pressure of the liquid in which the metal is immersed. In the case of liquid helium, a variation in vapour pressure of about 40 mm. of mercury corresponds to about 0·1° C.

Recently McLennan, Allen, and Wilhelm[20], in making a study of various alloys of the silver-tin, gold-tin, and gold-lead systems, found three outstanding features to characterize the results. First, in alloys with the super-conducting elements it was observed that gold and silver produced an effect on the transition temperature opposite to that produced by bismuth, antimony, and arsenic. When one observes

alloys containing the latter metals, one finds usually a pronounced elevation of the super-conducting temperature, while in alloys with gold and silver one finds an equally pronounced depression of that temperature. Secondly, it was noted that a binary alloy system composed of a super-conductor and a non-super-conductor does not necessarily have a unique transition temperature. Thirdly, it was found that with the alloy systems silver-tin, gold-tin, and gold-lead, the transition temperatures were higher for eutectic mixtures than for chemical compounds of the two metals constituting the alloys. The data compiled in table 3 will serve to illustrate these points. The element silver and the compound Ag_3Sn, it will be seen, have not been found to be super-conductors at any temperature reached up to the present.

It has been shown by Meissner[21] in an examination of the alloy systems In-Pb, In-Tl, Pb-Bi, Pb-Hg, Pb-Tl, Tl-Sn, and of other systems as well, that the super-

Table 3. Tin-silver alloys

Substance	Mixture	Percentage of tin	Transition temperature (°K.)
Sn	Pure	100	3·76
Ag + Sn	Eutectic alloy	96	3·52
Ag_3Sn + Sn	Eutectic alloy	50	3·57
Ag_3Sn + 3 per cent Sn	Mixture	30	2·3
Ag_3Sn	Compound	27	—
Ag	Pure	0·0	—

conducting transition temperature of an alloy is a continuous function of the concentration. From an extensive investigation of the thallium-tin system of alloys, J. F. Allen[22] has recently found some interesting features connected with the eutectic region. Both the transition temperature and the transition interval become a minimum for the eutectic. The transition interval increases from zero for pure tin to 1·3 degree for 78 per cent tin, and then decreases to 0·8 degree at the eutectic. The interval grows to 2·5 degree again near 40 per cent tin and again shrinks to 1·2 degree as the region of solid solution is approached. X-ray examination of these Tl-Sn alloys shows that both of the components possess maximum dimensions at the eutectic, and this indicates a close relation between super-conductivity and crystal structure.

From the work of Meissner, de Haas, and others, and from the X-ray measurements of Solomon and Morris Jones[23] on the Pb-Bi alloy system, a connexion appears to exist between the super-conducting transition temperature and the dimensions of the crystal lattice. There are indications that the electrostatic field of force of the lattice is probably a fundamental factor in determining whether a substance will become a super-conductor or not.

Mechanical strains and thermal dilatation. The application of mechanical stresses such as those of torsion and tension raises the transition temperature of a super-conducting metal, but observations made on the thermal dilatation of a lead rod showed no discontinuity when its temperature was lowered as it passed

through the transition temperature, 7·2° K., from the ordinarily conducting to the super-conducting state.

Action of a magnetic field. The application of a magnetic field delays the appearance of super-conductivity and causes it to appear in a metal at a lower temperature than normally.

If a metal in the super-conducting state be subjected to a gradually increasing magnetic field, a critical field strength is reached at which electrical resistance reappears in the metal. The strengths of the critical fields required for different

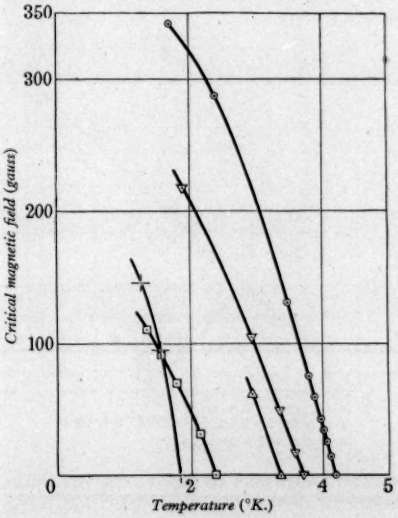


Figure 7.

⊙ Hg. ▽ Sn. △ In. ⊡ Tl. + Au_2Bi.

super-conductors vary; for example, an alloy of bismuth and lead at 1·91° K. requires a magnetic field of 26,700 gauss to restore the property of electrical resistance, while metallic tin at about the same temperature requires a field of only 220 gauss. In figure 7, curves are shown which give the characteristic values of the critical magnetic fields[24] as a function of temperature for a number of super-conducting metals.

Since the electrical resistance of super-conducting metals is zero, no heat is produced when electrical currents are passed through them. Currents of high intensity can therefore be passed through super-conducting wires of small diameter without melting them. Electric currents of more than 1000 amperes have been so

obtained in wires of small cross-section. The factor that imposes a limiting value upon the current strength is the magnetic field set up in the wire by the current itself. A critical value is reached when resistance is restored to the wire by the magnetic field.

Owing to the fact that metals in the super-conducting state have no electrical resistance, currents of electricity induced in rings of metals in this state will persist with undiminished intensity so long as the metals remain super-conducting. So far, it has been found impossible to detect with instruments of precision any diminution in the intensity of ring currents in super-conductors even after the lapse of a period so long as 13 hours.

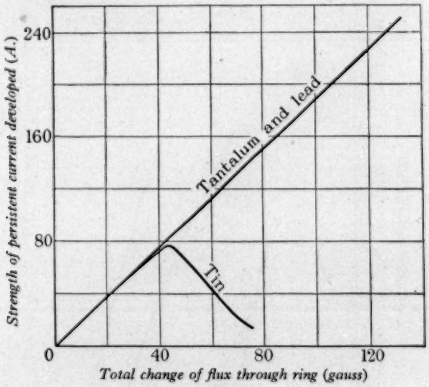


Figure 8.

Recently some experiments were carried out by McLennan, Allen, and Wilhelm[25] on the intensities of persistent currents of electricity induced in rings having the same dimensions and made of lead, of tin, and of tantalum, brought into the super-conducting state by the use of liquid helium. The currents in the rings were induced by the magnetic field provided by electric currents established in a circular coil of wire placed, in turn, coaxial with and close to each of the super-conducting rings. The results of these experiments are represented by the graph shown in figure 8. It was found that, for the weaker magnetic fields, equal changes of flux produced currents of equal magnitude in each of the three super-conductors. The magnitude of the persistent current developed depended not on the substance of the super-conducting ring but only on its dimensions and on the magnitude and form of the inducing magnetic field.

The case of tin is very interesting, since the values of the current in it agreed with those of the current in the others only up to fields of about 30 gauss. For inducing fields higher than this amount the strength of the persistent current

dropped off. Above this point, then, part of the ring must have been in a magnetic field, the strength of which had reached the critical value where resistance reappeared, i.e. an inner layer of the ring must have become non-super-conducting. As the field was increased above this point, one can suppose that the outside super-conducting portion of the ring became thinner and thinner until finally the whole ring became conducting only in the ordinary sense.

The fact that the same flux engenders the same persistent current in different super-conducting metals having the same size and form follows from an application

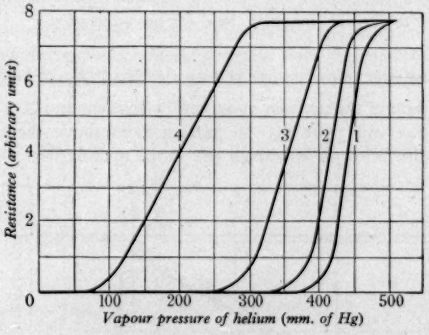


Figure 9. Lowering of transition temperature with increase of current.
(100 mm. pressure ≡ 2·64° K.; 250 mm. ≡ 3·22° K.; 500 mm. ≡ 3·79° K.)

Tin on constantan { Curve 1, 50 mA. d.-c.
Curve 2, 100 mA. d.-c.
Curve 3, 200 mA. d.-c.
Curve 4, 500 mA. d.-c.

of the equation $Ldi/dt = dB_A/dt$, or $i = B_A/L$. For rings of the same dimensions, the self-inductances would be identical; and in the super-conducting state the resistances of the three metals would be vanishingly small. Self-inductance and zero resistance are the two factors that make the magnitude of the induced current in different super-conductors the same.

Looking at the matter in another way, we see that the induced currents in the three super-conductors must be the same since the magnetic field of the persistent current must be equal in magnitude and distribution, but opposite in direction, to the flux of the exciting field.

In his experiments with mercury and with thin films of tin, Onnes found that an increase in the current through the specimen studied lowered the transition temperature. Curves which illustrate this point are shown in figure 9[26]. They were plotted from measurements made at Toronto on the conductivity of tin-coated wires at various temperatures. This lowering of the transition temperature, as already indicated, was probably due to the magnetic field set up by the current

through the specimen, since the application of a magnetic field to the region in which the specimen is placed has the effect of lowering the transition point. By agreement the accepted transition point is now generally defined as the temperature at which super-conductivity sets in when the current through the conductor is small.

Onnes was able to show that a finite resistance was restored to a specimen exhibiting super-conductivity when the current through it was raised to several hundred amperes. This point was cleared up when Silsbee[27] showed that the current which restored the resistance of the specimen studied produced a magnetic field in the wire equal to the critical magnetic field for the substance.

§ 7. SPECIFIC HEATS OF SUPER-CONDUCTING METALS

The specific heat of tin has been measured by Keesom and Kok[28] between the temperatures 3·741 and 3·505° K. In passing from above the transition point (3·7° K.) to a point below it, an abrupt rise in the specific heat was observed. It

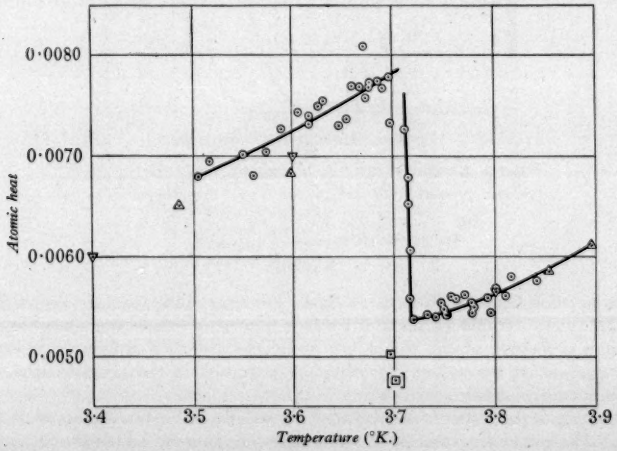


Figure 10. ⊙ Results of Keesom and Kok, *Leiden Comm.*, 221*e*. △ Results of Keesom and van den Ende, *Leiden Comm.*, 219*b*. ▽ Corrected results of Keesom and van den Ende. ⊡ Measurement made in presence of a stray magnetic field.

was found that the atomic heat increased from 0·0054 to 0·0078 calorie per degree K. The results of the measurements of the specific heat are shown in figure 10.

Observations were also made by the same investigators on the specific heats of thallium between 1·3° and 4·2° K. With this metal also an increase in the atomic heat amounting to 0·00148 calorie per degree K. was obtained in going from just

above the transition point (2·37° K.) to a point just below it. Recently Keesom and Kok made calorimetric experiments with thallium in connexion with the transitions in a number of constant magnetic fields from the super-conducting to the non-super-conducting state. A definite latent heat was measured for each field. The results seem to indicate that the transition from the super-conducting to the non-super-conducting state can be treated thermodynamically as a reversible process.

§ 8. HEAT CONDUCTION IN METALS AT LOW TEMPERATURES

Though the experiments are difficult to arrange and to carry out, de Haas and Bremmer[29] have succeeded at Leiden in measuring the conduction of heat at low temperatures by the super-conducting metals lead, tin, indium, and the lead-thallium alloy Pb-Tl_2. With all of them the thermal resistance fell to a minimum

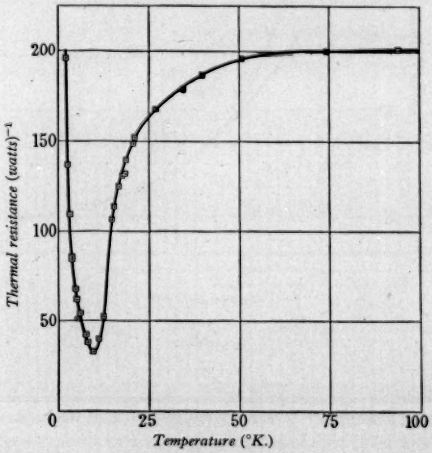


Figure 11.

value as the temperature was lowered. With lead and tin the temperature corresponding to the increase in thermal resistance was slightly over 9° K. With indium it was 18·3° K., and with Pb-Tl_2 it was a higher value still. In the super-conducting region for all of them the thermal resistance rapidly increased with decreasing temperatures. A characteristic curve illustrating the observations made with lead is shown in figure 11.

The observations made with tin, both without and with the application of a magnetic field, are represented by the curves in figure 12. From these it will be seen that in the super-conducting region, i.e. below 3·7° K., the thermal resistance was lowered when the magnetic field was applied. This would seem to show that

when a metal becomes super-conducting some constituent of it which is present at higher temperatures is no longer available for functioning in the conduction of heat by the metal.

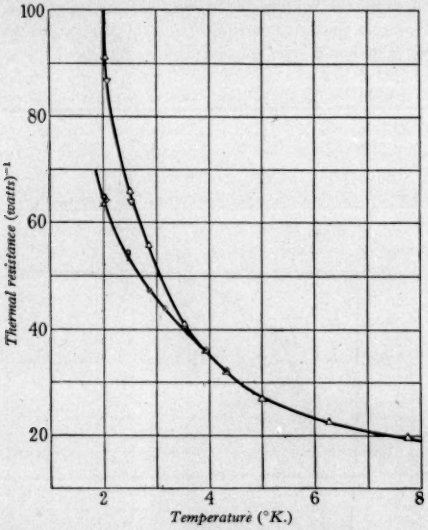


Figure 12. ♀ 192·6 gauss. ○ 257 gauss. ✧ 638 gauss. ♃ 428 gauss. ⊙ 244 gauss. △ 0 gauss. ▽ 74·9 gauss.

§ 9. MAGNETIC PROPERTIES OF SUPER-CONDUCTORS

Meissner and Ochsenfeld[30] recently discovered that when tin, which is feebly paramagnetic, and lead, which is feebly diamagnetic, were placed in a magnetic field of small intensity and gradually cooled to very low temperatures they suddenly at the transition point assumed zero magnetic permeability or a diamagnetic susceptibility of $-1/(4\pi)$. This discovery has been confirmed by Mendelssohn and Babbett[31] at Oxford and by Burton[32] and his collaborators at Toronto. In the Toronto experiments a hollow cylinder of tin was used as the super-conducting body. It is shown, with dimensions inscribed, in figure 13. The cylinder was placed with its axis perpendicular to the magnetic field. Coil no. 1 was wound round one side of the cylinder, coil no. 2 was placed inside the hollow cylinder, and coil no. 3 was placed near the surface of the cylinder with its plane perpendicular to the field. Coil no. 4 was wound so as to enclose the whole cylinder in the plane of the coil, and no. 5 was placed at the outer surface of the cylinder with its plane perpendicular to the applied magnetic field.

According to the results of Meissner and Ochsenfeld, which are illustrated by the configuration of the magnetic lines of force in figures 14 and 15, one would expect—when the temperature of the tin cylinder was taken through the transition point to a temperature definitely below the super-conducting temperature—to observe the following: in coil 1, a decrease to zero flux; in coil 2, no change in flux; in coil 3, an increase of twice the intensity of the applied field; in coil 4, a decrease in flux due to the wiping-out of the magnetic flux in tin; in coil 5, a decrease to zero

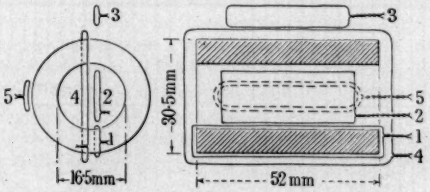


Figure 13.

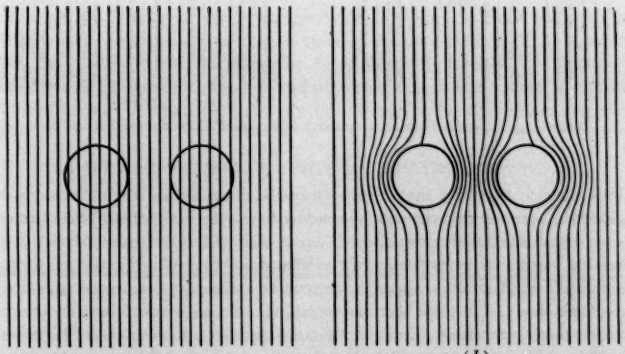


Figure 14. Parallel cylinders of tin in a magnetic field. (*a*) $T > 3{\cdot}7°$ K. (*b*) $T < 3{\cdot}7°$ K.

flux. The search coils in the Toronto experiments were arranged to be connected directly to a flux meter. The tin cylinder was passed from the super-conducting to the non-super-conducting state as desired by simply altering the vapour pressure of the liquid helium in which the tin cylinder was immersed.

It was found that as the tin cylinder was taken from above the transition temperature to below it, coil 1 showed a decrease of 90 per cent in the flux; coil 2 showed a slight increase up to 10 per cent; coil 3 showed an increase of 25 per cent; coil 4 a decrease of 30 per cent; and coil 5 a decrease of from 20 to 25 per cent.

It should be noted that coil 3 projected out about 5 mm. from the surface of the cylinder, where there was undoubtedly a magnetic field of high gradient, and also that coil 5 of necessity enclosed a considerable space where the field was not theoretically reduced to zero but only relatively weakened. The field strengths used were approximately 30, 150 and 200 gauss.

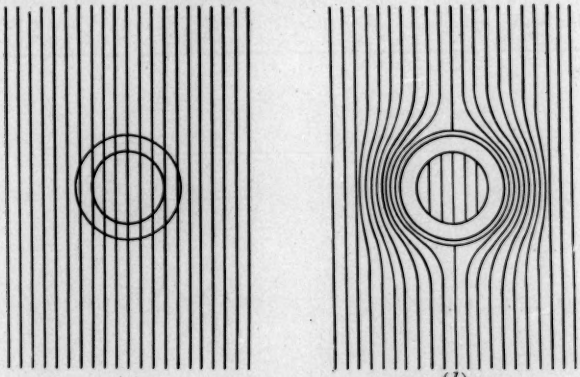


Figure 15. Cylinder of tin in a magnetic field. (*a*) $T > 3{\cdot}7°$ K. (*b*) $T < 3{\cdot}7°$ K.

§ 10. NEW MAGNETIC TEST FOR SUPER-CONDUCTIVITY

It is at once apparent that Meissner and Ochsenfeld by their discovery have provided a magnetic effect that can be used to make a decisive test for the appearance of super-conductivity. Burton and his associates at Toronto have applied it with interesting results in a new study of the alloy known as Wood's metal. As far as former super-conductivity measurements on this alloy are concerned one element alone of the alloy might have been responsible for the super-conductivity observed, with the other elements remaining in the non-super-conducting state. If, however, on applying the new test it were found that the magnetic field disappeared for the complete volume of the alloy, it would mean that the alloy as a whole became super-conducting. This complete disappearance of the flux was actually observed with Wood's alloy. The exploring coils, it may be added, were embedded in the molten metal (melting point 65° C.) and remained there when the alloy was cooled to its transition temperature. Wood's alloy, it is well known, has a constitution of 50 per cent bismuth, 25 per cent lead, 12·5 per cent tin, and 12·5 per cent cadmium.

The method is also being applied to a study of alloys that contain a number of phase constituents, with a view to determining the temperatures at which each of the different phases become super-conducting.

In another experiment carried out by Burton[33] at Toronto the magnetic-field

test was used to examine the super-conductivity of an emulsion of mercury and lard. A microscopic examination of the emulsion indicated a surprising uniformity in the size of the globules of mercury (of the order of 10^{-5} cm.). The emulsion was found to be quite non-conducting electrically at ordinary temperatures. At the lowest temperatures three tests were made, namely (1) with pure mercury, (2) with the mercury emulsion, and (3) with pure lard.

An exploring coil was embedded in each of the matrices at room temperature, so that when the alloy cooled down and was frozen the coil was completely embedded. The amount of mercury in the emulsion was almost 50 per cent by volume. When the three substances were cooled down in a constant magnetic field the following results were obtained: (1) Magnetic flux completely disappeared as the temperature passed the mercury transition point. (2) A reduction in the flux of between 30 and 40 per cent. (3) No change in the flux.

These results showed that mercury can exhibit the super-conducting property even when it is in the form of very small globules 10^{-5} cm. in diameter.

§ 11. SUPER-CONDUCTIVITY WITH RAPIDLY ALTERNATING ELECTRIC FIELDS

The results of some of the experiments on super-conductivity suggest that an orientation or re-orientation of one or more of the constituents of the conducting metals takes place when the transition temperature is reached. As times of relaxation are generally associated with orientation effects, it was natural to investigate the phenomenon of super-conductivity by the use of rapidly alternating electric fields.

§ 12. ABSORPTION OF β-RAYS[34]

In one of these researches the absorption of β-rays by a thin sheet of lead was investigated when the lead was gradually cooled from a temperature a few degrees above to a few degrees below the critical transition temperature of 7·2° K. The β-rays used were those emitted by mesothorium, and the lead sheet had a thickness sufficient to absorb, at ordinary room temperature, 50 per cent of the β-rays issuing from the mesothorium. In these experiments no measurable variation or discontinuity was detected in the absorption coefficient as the temperature of the lead was lowered through the critical value (7·2° K.). The high-velocity electrons from the mesothorium apparently encountered just as much resistance in their passage through the lead with the latter in the super-conducting state as when the lead possessed the normal conductivity exhibited at the higher temperatures. This investigation gave definite proof that although resistance in the super-conducting state is zero for currents carried by slow-moving electrons, it is not zero but maintains a normal value for currents carried by high-speed electrons.

Looking at the matter in another way, this result indicates, if the de Broglie wave equation $\lambda = h/(mv)$ applies, that lead at the lowest temperatures cannot exhibit super-conductivity when subjected to alternating electric fields with frequencies of the order of 10^{21} cycles per sec.

§ 13. PHOTO-ELECTRIC AND LIGHT-ABSORPTION EXPERIMENTS

In a second series of experiments thin films of lead were deposited on plates of glas s and of quartz, sometimes by cathode spluttering and at other times by vaporization of metallic lead. These films were subjected to a series of decreasing temperatures, commencing a few degrees above 7·2° K. and ending at the temperature of liquid helium (4·2° K.). The photo-electric effect and the absorption of visible light were in turn investigated with these films, and measurements were taken approximately by steps of a fraction of a degree as the temperatures of the films were lowered. In these experiments no measurable discontinuity was observed in the results of the measurements on the photo-electric effect, nor in the results of those on the coefficient of absorption of the light waves when the lead films traversed were passed through the transition temperature of 7·2° K. These results were therefore taken to indicate that super-conductivity with lead is a phenomenon that cannot be exhibited when electric fields alternating with a frequency approximately equal to or greater than 10^{14} per sec. are used.

It is clear, however, since super-conductivity can be brought into evidence by the use of unidirectional fields, i.e. with fields of zero frequency, that there must exist some critical alternating field, with a frequency between zero and 10^{14} per sec., by the use of which super-conductivity should just be detectable.

§ 14. EXPERIMENTS WITH RADIO-FREQUENCY ELECTRIC FIELDS[35]

Through the development which has taken place in recent times, it is a comparatively simple matter to arrange combinations of oscillating valve systems capable of providing alternating electric fields with frequencies as high as 10^7 or even 10^8 cycles per sec. Some experiments were therefore made with radio-frequency fields of approximately 10^7 cycles per sec., corresponding to a wave-length of about 30 metres. The experiments and apparatus used, together with the theory applicable, have been fully described elsewhere and are therefore only briefly referred to here.

A resonant circuit was constructed consisting of a coil of the wire under investigation, which was wound on a former in series with a condenser mounted on the same former. The entire electrical circuit was made of the metal under investigation. The resonators used in the experiments were made of lead, of tin, and of tantalum. The plates of the condenser were kept at the proper distance apart by fibre washers on the central spindle. This resonator, in the experiments, was kept in the Dewar (unsilvered) flask in which the liquid helium or gaseous helium of a low temperature was produced. Temperatures were measured with helium-gas thermometers or were deduced from the vapour pressure of the liquid helium.

A reactance method was used and the plate and grid coils of the oscillating system were placed, properly screened, external to the vacuum flask containing the resonant circuit and immediately below it. The magnitudes of the currents induced in the resonance coil were measured by their reaction upon the oscillations in the

generating circuit as indicated by the plate current of the latter. Measurements made in this way enabled one to deduce the value of the a.-c. resistance in the resonant circuit when fields of different frequencies were used at various selected low temperatures.

The results obtained are typified by those for a solid tantalum wire illustrated by the curves in figure 16. These curves indicate that the transition point was lower where fields with a frequency of $11{\cdot}4 \times 10^8$ cycles per sec. were used than it was with unidirectional fields. The curious hump[36] in the high-frequency curve was originally ascribed to an experimental error. A closer study, however, has seemed to show that with alternating applied fields there is a phase variation in the current as one passes from the surface to the centre of a solid wire which may account for the

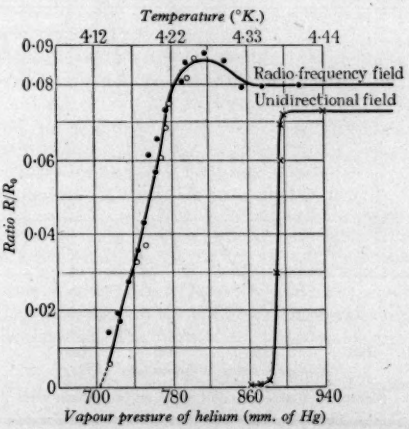


Figure 16. O Points obtained with a very weak coupling.

effect. A difference of 180° or more between the phase of the current at the surface and that at the centre may exist. It may happen then that a decrease in the ohmic resistance of the central elements of a conductor produced by a lowering of the temperature will result in an increase in the total a.-c. resistance of the conducting wire, through the current being confined to a thinner shell.

The initial experiments with alternating electric fields applied to solid conducting wires seemed to show that at frequencies higher than 10^9 cycles per sec. super-conductivity was not obtainable with a wire of any kind of metal. Later experiments, however, made with looser coupling between the oscillating and resonance circuits, have yielded results, particularly with thin films of metals used in place of solid wires, that prove this conclusion to be not valid generally. Figure 17 contains curves that illustrate the results obtained with a thin (0·002 mm.) film of

tin "wiped" upon a constantan wire having a diameter of 0·016 cm. Curve 1 represents the d.-c. resistance measurements with a current of 50 mA., curve 2 gives the high-frequency resistance with an alternating current of 25 mA. having a frequency of 10^7 cycles per sec., and curve 3 shows the high-frequency resistances obtained with currents of varying value but always having a frequency of 3×10^7 cycles per sec.

Three striking characteristics of super-conductors are revealed by these curves: (*a*) The transition temperature is the same for both unidirectional and high-

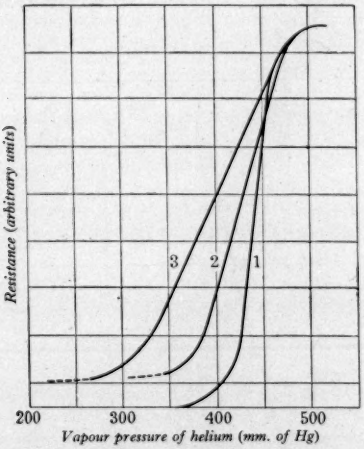


Figure 17. Curves for film of tin on constantan wire.
(250 mm. pressure ≡ 3·22° K.; 500 mm. ≡ 3·79° K.)
Curve 1. Direct current, 50 mA.
Curve 2. 10^7 cycles per sec., 25 mA.
Curve 3. 3×10^7 cycles per sec., varying current.

frequency fields when small current values are used. (*b*) The disappearance of the resistance, or the vanishing point, is at a lower temperature for high-frequency resistance than for d.-c. resistance: in fact, the value of the apparent resistance with the high-frequency currents never becomes absolutely zero as the d.-c. resistance does. (This is probably because the apparent resistance is partly ohmic and partly induced.) (*c*) There is no indication of a slight rise in resistance before the films become super-conducting to high-frequency currents, such as is invariably observed with solid wires of the same material.

It is interesting to note that with films there is no appreciable change in the transition temperature as between direct currents, and alternating currents for frequencies up to 3×10^7 cycles per sec.

In view of the results obtained with films of lead in a study of the absorption of β-rays, the transmission of light, and the photo-electric effect, which showed no discontinuities at the transition temperature for direct currents, it would seem desirable to pursue the investigation further with currents having frequencies between 3×10^7 and 10^{14} cycles per sec.

§ 15. THE EFFECT OF SUPERIMPOSING DIRECT CURRENT AND ALTERNATING CURRENTS

Experiments have been carried out at Toronto with both solid wires and thin films of super-conducting metals—tin, for example—to determine (1) the effect of superimposing direct current on the high-frequency current while measuring the high-frequency resistance, and (2) the effect of superimposing high-frequency current on the direct current when measuring the d.-c. resistance.

The results obtained with films of tin deposited on constantan wires illustrate the effects observed: (1) There is no appreciable effect on the transition point for high-frequency resistance due to the superposition of a direct current for small values of the latter. (2) The effect of superimposing high-frequency currents on direct currents while measuring d.-c. resistance is to lower the transition or critical temperatures of the film investigated. The effect depends on the magnitude of the alternating current, and not at all on its frequency.

§ 16. FILM THICKNESS AND TRANSITION TEMPERATURES

Some experiments recently made by Burton[37] and his co-workers at Toronto on transition temperatures for films of tin of different thicknesses have yielded results of considerable importance. The films were made by electro-plating wires of

Table 4. Transition points for films of different thicknesses.

Thickness of tin film (cm.)	Material of wire core	Diameter of core (cm.)	Transition point (°K.) for currents of:		
			6 mA.	18 mA.	200 mA.
200×10^{-5}	Constantan	0·056	—	—	3·78
90×10^{-5}	Constantan	0·056	3·78	3·76	3·74
37×10^{-5}	Constantan	0·056	3·77	3·76	3·73
10×10^{-5}	Constantan	0·056	3·71	3·66	3·61
9×10^{-5}	Phosphor bronze	0·040	—	3·64	3·54
$6{\cdot}8 \times 10^{-5}$	Constantan	0·056	*	3·60	3·49
5×10^{-5}	Constantan	0·056	—	2·54	—
4×10^{-5}	Chromel	0·040	—	3·25	3·00
4×10^{-5}	Manganin	0·045	—	3·12	2·92
$2 \times 10^{-5} - 3 \times 10^{-5}$	Constantan	0·056	3·23	3·10	2·48
$2{\cdot}5 \times 10^{-5}$	Constantan	0·056	Not super-conducting above 2° K.		
$1{\cdot}6 \times 10^{-5}$	Constantan	0·056	,,		
$0{\cdot}8 \times 10^{-5}$	Constantan	0·056	,,		

* 3·64° K. at 2 mA.

Table 5. The effect of outer coating on the transition point.

Thickness of tin film (cm.)	Outer coating		Transition point (°K.) for currents of:			
	Material	Thickness (cm.)	6 mA.	18 mA.	200 mA.	500 mA.
200×10^{-5}	—	—	—	—	3·78	—
200×10^{-5}	Cu	80×10^{-5}	—	—	3·75	3·74
90×10^{-5}	—	—	3·78	3·76	3·74	3·70
90×10^{-5}	Ni	86×10^{-5}	3·71	3·69	3·65	3·58
90×10^{-5}	Cu	80×10^{-5}	3·71	3·69	3·65	3·58
50×10^{-5}	—	—	—	3·75	3·72	3·69
50×10^{-5}	Ni	241×10^{-5}	—	3·70	3·64	3·54
50×10^{-5}	Cu	7×10^{-5}	3·72	3·69	3·64	—
37×10^{-5}	Cu	37×10^{-5}	Not super-conducting above 2° K.			
20×10^{-5}	Cu	10×10^{-5}	,,			
18×10^{-5}	Cu	40×10^{-5}	,,			
15×10^{-5}	Ni	30×10^{-5}	,,			
9×10^{-5}	Cu	100×10^{-5}	,,			

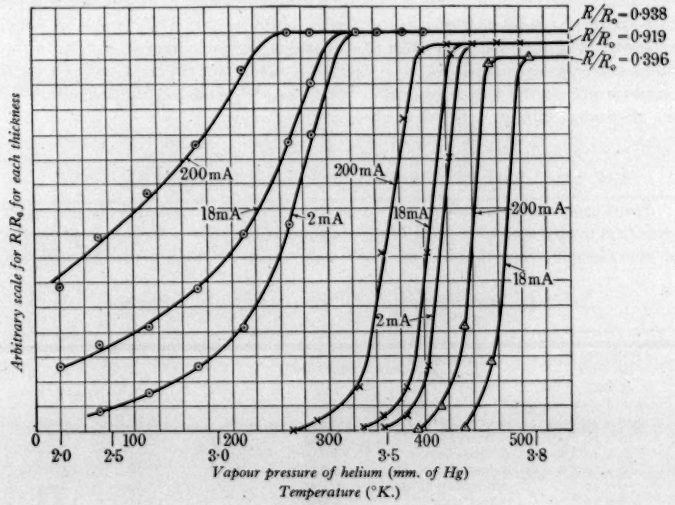


Figure 18.

⊙ Thickness of tin film, 2×10^{-5} to 3×10^{-5} cm.
× Thickness of tin film, 6×10^{-5} to 8×10^{-5} cm.
△ Thickness of tin film, 10×10^{-5} cm.

different metals according to the method of Mathers and Bell[38], and the transition temperatures were determined with each film for currents of 6, 18 and 200 mA.

A typical set of results is given in table 4 and illustrated diagrammatically in

figure 18. It will be seen that the transition point of a tin film recurred at a lower and lower temperature as the film was reduced below 10×10^{-5} cm. Above this thickness the transition temperature was practically constant for currents of small intensity. For thicknesses of 2×10^{-5} to 3×10^{-5} cm. the films did not become super-conducting at all, even when temperatures as low as 2° K. were used.

A striking feature shown by the results in table 4 is that the influence of increasing currents in lowering the transition temperature is very much more marked with

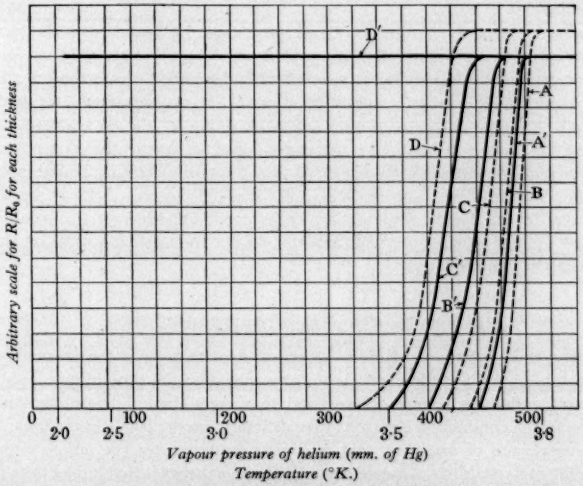


Figure 19. Curves for film of tin on constantan core, with and without coating of copper. Effect of (i) thickness of film, (ii) surface coating.

---- Uncoated tin films. —— Coated with copper.

For curves A and A', thickness of tin = 200×10^{-5} cm.
For curves B and B', thickness of tin = 90×10^{-5} cm.
For curves C and C', thickness of tin = 50×10^{-5} cm.
For curves D and D', thickness of tin = $6{\cdot}8 \times 10^{-5}$ cm.
D' not super-conducting above 2° K.
Thickness of copper coating > 10^{-4} cm.
Current, 18 mA.

thin films than with thick ones. This would indicate that it may be advisable to consider current density in a super-conductor rather than current strength as the more important factor in developing a theory for super-conductivity. In these experiments with tin deposited on wire cores Burton also examined the effect of coating the thin films of tin with an outer film of copper or nickel. The very striking result obtained was that a tin film which ordinarily became super-conducting near its normal transition point (say, for example, 3·7° K.) lost its super-conducting property when covered with a film of copper or nickel. It was found that neither the thickness of this outer coating (from 7×10^{-5} cm. to 10^{-3} cm.) nor the material (nickel or copper) varied the depression of the transition temperature. Systematic

investigation of films of various thicknesses showed that there was a limiting thickness of tin (about 40×10^{-5} cm.) below which this loss of super-conductivity showed itself; but even for much thicker tin films the transition point was lowered when the outer coating was put on. The results are given in table 5, and curves representing them are shown in figure 19. It will be seen that the coated tin film of thickness 50×10^{-5} cm. became super-conducting at about 3° K., while that of thickness 37×10^{-5} cm. was prevented from becoming so—certainly for a temperature as low as 2° K.

These experiments on thin films, and those made by Burton with the mercury-lard emulsion, are among the most interesting as yet carried out in the investigation of super-conduction in metals. It seems highly desirable to have those made with films repeated at still lower temperatures than 2° K., possibly at the temperatures obtainable through adiabatic cooling brought about by the demagnetization of paramagnetic salts. It would be interesting also to repeat the experiments of Meissner and Ochsenfeld with deposited thin films of different thicknesses, both when uncoated and when coated with outer films of copper. A study, by the use of X rays, of the crystal structure of these deposited films of tin—coated and uncoated—might possibly too yield results of value in connexion with the elucidation of super-conductivity.

§ 17. THEORETICAL CONSIDERATIONS

Among those who have offered theoretical explanations of the phenomenon of super-conductivity is Brillouin[39], who has considered the possibility of assuming the super-conducting state to correspond to a meta-stable current-carrying distribution of the conduction electrons over different energy levels in the ionic lattice of the metal, and of assuming that at higher temperatures the lattice vibrations destroy the meta-stability. One defect in Brillouin's theory is that it does not provide for the sharp transition to super-conductivity observed experimentally.

A theory put forward by R. de L. Kronig[40] seems better suited to explain the facts. On this theory a part of the conduction electrons are supposed to form below the transition point 1-dimensional chains which can move freely through the ionic lattice. The transition temperature can be considered to be the melting point of this electronic lattice.

Amplifying Kronig's theory, Frenkel[41] has recently pointed out that 1-dimensional lattices, or rather movable chains of bound electrons, will be stable and will suffice for the explanation of super-conductivity.

Such chains need not be movable in all directions. It will suffice to suppose them to be movable in one particular crystallographic direction corresponding to the smallest spacing between the atoms, the dielectric constant being infinite for this direction and preserving a finite value for all the others. It is customary to describe the super-conducting state of a metal by setting its specific conductivity equal to infinity. Frenkel directs attention to another possibility, namely, that the super-conducting state can be described much more adequately by setting equal to

infinity the dielectric constant (κ) of the substance, its conductivity (σ) remaining finite or even becoming equal to zero. With Frenkel's modification of Kronig's theory a super-conductor could be described as a body with zero magnetic permeability ($\mu = 0$) and a dielectric constant given by $\kappa = \infty$, its electrical conductivity (σ) in the exact sense of the term being finite or even zero.

Frenkel makes a comparison of the mechanism of ordinary electric conduction (σ finite) and of ordinary polarization (κ finite). In the former case the electrons called "free" move independently, the conduction current being constituted by a drift motion due to the action of an external electric field and superimposed on the unperturbed random motion of the individual electrons. In the latter case the electrons called "bound" are displaced by the electric field simultaneously in the same direction, the polarization current being due to an orderly collective motion of all the electrons. Under normal conditions the displacement of the electrons with regard to the respective atoms remains small compared with the interatomic distances; this corresponds to a finite value of the dielectric constant. The assumption that the latter becomes infinite means that under the action of an infinitesimal field the electrons are displaced simultaneously over finite distances, each of them passing successively from one atom to the next, like a chain gliding over a toothed track.

According to Frenkel a collective motion of the "bound" electrons will constitute an electric current just as much as the individual motion of the free electrons, but a polarization current rather than a conduction one. The electrostatic mutual action of the electrons moving collectively in a chain-like way will stabilize them against the perturbing action of the heat motion of the crystal lattice. This will result in the permanence of the polarization current after the disappearance of the electric field by which it was started. This permanence, which has been interpreted hitherto as corresponding to an infinite value of the specific conductivity, must be interpreted on the new view as corresponding to an infinite value of the dielectric constant.

Though the explanation of super-conductivity offered by Frenkel is essentially of a formal character, it should be mentioned that he has given strong reasons in support of the view that with a metal one should expect to realize a physical state characterized by an infinite dielectric constant. Grayson Smith of Toronto[42], who has recently made a critical review of all available experimental data on the super-conducting state of metals, concludes that the only way of accounting for the results of the Meissner and Ochsenfeld experiments is to assume that there arises spontaneously, as the metal passes through the transition point, a system of local electric currents giving a sort of microstructure to the metal. Such a conception had already been suggested by Frenkel, but Grayson Smith has pointed out that the magnetic property of super-conductors allows an estimate to be made of the minimum breadth of these vortex systems. It appears that they must be greater than 10^{-5} cm. in breadth. This, it will be noted, is in fair agreement with the actual thickness of thin films below which there is a marked interference with the onset of super-conductivity.

In concluding this lecture I should like to draw attention to the attempt now being made by Kapitza in the Royal Society-Mond Laboratory at Cambridge to liquefy helium without having recourse to the use of liquid hydrogen as an intermediary cooling agent. Kapitza proposes to use liquid air alone for external cooling and to reduce the temperature of the gaseous helium below its inversion point by making the gas do mechanical work against a small engine of special design included in the helium gas cycle. From present indications it is highly probable that Kapitza's attempt will culminate in a brilliant success.

REFERENCES

(1) W. H. KEESOM. *Nature*, Lond., **118**, 81 (1926); also *Leiden Comm.* No. 184 (*b*).
(2) J. DEWAR. "Liquefaction of gases," *Ency. Brit.* (11th ed.).
(3) W. H. KEESOM. *Proc. K. Akad. wet. Amst.* **35**, No. 2, 1 (1932); also *Leiden Comm.* No. 219 (1932).
(4) W. H. KEESOM. *Leiden Comm.* No. 202 (*b*).
(5) W. H. KEESOM and A. P. KEESOM. *Leiden Comm. Supplement*, No. 76 (*b*).
(6) W. H. KEESOM. *Leiden Comm. Supplement*, No. 75 (*a*).
(7) W. H. KEESOM and K. CLUSIUS. *Leiden Comm.* No. 216 (1931).
(8) W. H. KEESOM and K. CLUSIUS. *Leiden Comm.* No. 219 (*f*), p. 60 (1932).
(9) W. H. KEESOM and A. P. KEESOM. *Leiden Comm.* No. 221 (*d*) (1932).
W. H. KEESOM and K. CLUSIUS. *Leiden Comm.* No. 219 (*e*).
(10) F. SIMON. *Z. phys. Chem.* **2**, No. 4 (1927), **6**, No. 1 (1929), **6**, No. 5 (1930).
W. H. KEESOM and J. H. C. LISMAN. *Leiden Comm.* No. 213 (*e*), p. 43 (1931), No. 221 (*a*), p. 1 (1932).
(11) R. KAISCHEW and F. SIMON. *Nature*, Lond., **133**, 460 (1934).
(12) F. A. LINDEMANN and T. C. KEELEY. *Nature*, Lond., **131**, 191 (1933).
(13) P. DEBYE. *Ann. Phys.*, Lpz., **81**, 1154 (1926).
(14) W. F. GIAUQUE. *J. Amer. chem. Soc.* **49**, 1864, 1870 (1927).
(15) W. F. GIAUQUE and C. W. CLARKE. *J. Amer. chem. Soc.* **54**, 3135 (1932).
(16) W. J. DE HAAS, E. C. WIERSMA and H. A. KRAMERS. *Naturwissenschaften*, **21**, 467 (1933); also *Nature*, Lond., **132**, 372 (1933).
(17) N. KURTI and F. SIMON. *Nature*, Lond., **133**, 907 (1934).
(18) J. DEWAR and J. A. FLEMING. *Phil. Mag.* **35**, 271 (1893).
(19) W. J. DE HAAS. *Leiden Comm.* Nos. 209 (*a*), 199 (*c*) and 197 (*a*); also *Naturwissenschaften*, **19**, 106 (1931).
W. J. DE HAAS and J. VOOGD. *Int. elect. Congr.* Report No. 10, Paris, Section 1, p. 19 (1932).
(20) J. C. MCLENNAN, J. F. ALLEN and J. O. WILHELM. *Phil. Mag.* **13**, 1196 (1932).
(21) W. MEISSNER. *Z. Phys.* **58**, 570 (1929), **65**, 30 (1930).
W. J. DE HAAS and J. VOOGD. *Leiden Comm. Supplement*, p. 73 (1932).
(22) J. F. ALLEN. *Phil. Mag.* ser. 7, **16**, 1005 (1933).
(23) D. SOLOMON and W. MORRIS JONES. *Phil. Mag.* **10**, 470 (1930), **11**, 1090 (1931).
(24) W. J. DE HAAS and J. VOOGD. *Int. elect. Congr.* Report No. 10, Paris, Section 1, p. 19 (1932).
(25) J. C. MCLENNAN, J. F. ALLEN and J. O. WILHELM. *Phil. Mag.* ser. 7, **14**, 168 (1932).
(26) E. F. BURTON. *The Phenomena of Superconductivity*, p. 53 (1934) (University of Toronto Press and Oxford University Press).
(27) F. B. SILSBEE. *J. Wash. Acad. Sci.* **6**, 597 (1916); *Sci. Pap. U.S. Bur. Stand.* No. 14, p. 307 (1917), No. 22, p. 556 (1927).
(28) W. H. KEESOM and J. A. KOK. *Leiden Comm.* No. 221 (*e*) (1932), Nos. 230 (*c*), 230 (*e*) (1934).

(29) W. J. DE HAAS and H. BREMMER. *Leiden Comm.* No. 214 (*d*) (1931).
(30) W. MEISSNER and R. OCHSENFELD. *Naturwissenschaften*, **21**, 787 (1933).
(31) K. MENDELSSOHN and J. D. BABBETT. *Nature*, Lond., **133**, 459 (1934).
(32) E. F. BURTON. *Nature*, Lond., **133**, 684 (1934).
(33) From a private communication from Prof. Burton to the author.
(34) J. C. MCLENNAN, J. H. MCLEOD and J. O. WILHELM. *Trans. roy. Soc. Can.* **23**, 3269 (1930).
J. C. MCLENNAN, J. H. MCLEOD and R. G. HUNTER. *Trans. roy. Soc. Can.* **24**, 31 (1930).
(35) J. C. MCLENNAN, A. C. BURTON, A. PITT and J. O. WILHELM. *Phil. Mag.* **12**, 707 (1931); also *Proc. roy. Soc.* A, **136**, 52 (1932), **138**, 245 (1932).
(36) E. F. BURTON. *The Phenomena of Superconductivity*, p. 69.
(37) E. F. BURTON and others. *Nature*, Lond., **133**, 459 (1934).
(38) F. C. MATHERS and W. H. BELL. *Trans. Amer. electrochem. Soc.* **38**, 135 (1920).
(39) L. BRILLOUIN. *J. Phys. Radium*, **4**, 333, 677 (1933).
(40) R. DE L. KRONIG. *Z. Phys.* **78**, 744 (1932), **80**, 203 (1933).
(41) J. FRENKEL. *Nature*, Lond., **133**, 730 (1934).
(42) From a private preliminary communication.

X RAYS

By THE X-RAY STAFF, PHYSICS DEPARTMENT,
National Physical Laboratory

General Editor: G. W. C. Kaye, O.B.E., M.A., D.Sc., A.R.C.S., F.Inst.P.

Authors of Sections

§ 1. Production of X rays, by G. E. Bell, B.Sc., A.Inst.P.
§ 2. Ionization in gases, by I. Backhurst, M.Sc.
§ 3. Absorption of X rays, by I. Backhurst, M.Sc.
§ 4. Scattering of X rays, by I. Backhurst, M.Sc.
§ 5. Measurement of intensity, quality, etc., by W. Binks, M.Sc., and G. E. Bell, B.Sc., A.Inst.P.
§ 6. X-ray spectra, by G. Shearer, M.A., D.Sc., Ph.D., F.Inst.P. With a sub-section on continuous spectrum by G. E. Bell, B.Sc., A.Inst.P.
§ 7. X-ray optics, by G. Shearer, M.A., D.Sc., Ph.D., F.Inst.P.
§ 8. Applications of X rays, by W. A. Wood.

§ 1. PRODUCTION OF X RAYS

High tension generators

In the early days of X-ray work there was no practical alternative to the induction coil fitted with some form of interrupter, but despite a great deal of research and considerable improvements in coil and interrupter design[1] the steadiness of the potential generated left much to be desired, while the current output was very limited. In 1908 high-tension transformers giving suitable voltages and of convenient dimensions were introduced and about 1915 hot-cathode rectifying valves[2] capable of suppressing high inverse voltages were developed. Almost immediately Hull[3] showed that by the use of a transformer, rectifying valves and suitable condensers, high constant voltages could be conveniently produced. Various arrangements of rectifiers and condensers have been described[4] whereby the output voltage of a generator can be made many times that of the high-tension transformer and generators of this type capable of giving constant voltages up to a million volts have been made for laboratory use[5].

Constant-voltage generators of the condenser type are comparatively inexpensive, and very suitable for giving X-ray tube currents up to about 20 mA. When high tube currents at approximately constant voltage are required it is usual to have recourse to a three-phase high-tension transformer, full-wave rectification being effected by six valves. Generators of this type have been built for outputs of 2000 mA. at about 100 kV. (peak). The overall voltage variation of a three-phase generator is about ± 7 per cent from the mean voltage. Three-phase generators capable of giving voltages much over 100 kV. are somewhat expensive.

A few workers have used high-frequency (Tesla) transformers[6] with some measure of success. The high voltages are produced in this manner by fairly inexpensive apparatus, but the high frequency of alternation causes certain technical difficulties while the current output is rather low. An alternative inexpensive source of very high voltages is the impulse generator[7]. Apparatus of this type has been constructed for upwards of two million volts[8] but the intermittent action of the generator combined with some uncertainty as to the voltage of the individual discharges render it unsuitable for many types of quantitative work.

The rectifiers used in high voltage generators are almost invariably either of the hot-cathode type or the mechanical type. Rectifiers of the hot-cathode type have the advantages of silent and corona-free operation, and are capable of passing heavy currents with comparatively small loss in voltage. Recently a form of mercury vapour rectifier has been introduced in which the voltage drop is only 40 volts for a current of 1000 mA. Mechanical rectifiers are in general somewhat unsteady in operation and have the additional disadvantages of being noisy and producing nitrous fumes. A generator employing an improved form of mechanical rectifier and capable of giving voltages up to 300 kV. (peak) has been described by Sarsfield[9].

X-ray tubes

From the discovery of X rays until 1913 the only tubes used were of the so-called "gas" type, in which the necessary electrons are produced by the ionization of the residual air in the tube. Such tubes have been developed to a high degree of efficiency, but suffer from several defects, the chief of which is that in general the tube current and exciting voltage cannot be varied independently. Some difficulty is experienced in maintaining the pressure of the residual air at the required value, since the vacuum tends to improve as the tube is operated, but this may be largely overcome by the use of automatic regulators. In 1913 Coolidge developed the high-vacuum hot-cathode tube[10], in which the source of electrons is a heated tungsten filament. Tubes of this type behave as ordinary diodes, and the electron current saturates at a comparatively low voltage, so that over a wide range, current and voltage may be varied independently of each other. Hot-cathode tubes in general operate more steadily than gas tubes, which they have now replaced for most purposes.

For crystal analysis and similar work where a high degree of steadiness in operation is not required and simplicity of construction is desirable, demountable gas X-ray tubes continuously evacuated by suitable pumps are extensively used.

About 1920 attention was first called to the dangers arising out of exposure to X radiation (see p. 238) and in 1924 Bouwers[11] described an X-ray tube so constructed that except for the useful beam no radiation escaped from the tube. Further improvements in the safety of X-ray equipment have been secured by the development of shock-proof tubes which may be connected to the high-tension supply by means of shock-proof cables. Several different forms of shock-proof tubes[12,13] have been described, but the general principle underlying the construction of most of them is the surrounding of the tube by an earthed metal casing. The chief

practical difficulties lie in making the casing conveniently small and equalizing the electrical stresses in the envelope of the tube. In some forms of tube these difficulties are met by the correct proportioning and shaping of the high-potential ends of the tube and the case while in others advantage is taken of insulating sheaths or oil within the earthed case. Shock-proof tubes have been constructed for operation at voltages up to 250 kV. peak and capable of passing currents up to 10 mA. when cooled by circulating water or oil.

Allibone and his collaborators[14] have described a self-protected shock-proof X-ray tube for hospital use which is continuously evacuated by means of oil-vapour diffusion pumps. The tube is operated by a constant-voltage generator embodying continuously evacuated rectifying valves, and the electrical design of the system is such that the process of evacuation is entirely automatic.

Recently some advance has been made in the construction of X-ray tubes for very high voltages. Allibone described a tube capable of operating at 500 kV. (Tesla excitation) when oil-immersed, while Lauritsen[15] has constructed, for medical purposes, an X-ray tube operating at 1100 kV. (peak). The latter tube consists of a central metal discharge chamber some seven feet long, to each end of which are attached insulating glass cylinders totalling about eleven feet in length. The tube is of the hot-cathode type, and continuously evacuated by means of mercury vapour diffusion pumps. For protection against stray X rays, the discharge chamber is provided with a closely fitting sleeve of lead sheet 2 inches thick, while the electrical safety of patients is assured since the earthed discharge chamber is the only part of the tube in the treatment room.

§ 2. IONIZATION IN GASES

The passage of an electric charge through air was the subject of an experiment by Coulomb as early as 1785 and since then a great wealth of data has been compiled concerning the many and important phenomena connected with gaseous ionization[1]. It has been established that the ion, or charge carrier in a gas, may be an electron, a charged molecule or small group of molecules, or a charged particle of a solid or liquid impurity, while any particular ion during its passage through a gas may change its type due to collision with gas molecules or dust.

Ionization may be brought about in many ways, as by the action of ultra-violet, X, α, β, γ, or cosmic rays, flames or other chemical reaction, heated solids, or by bubbling gas through liquids or splashing. The bombardment of molecules by electrons or positive ions results in ionization when appropriate critical energy values are attained, and has formed the basis of numerous experimental studies of ionization and resonance potentials.

The speed with which an ion travels per unit potential gradient, or its "mobility," is influenced by many factors, some of which are not yet too well understood. The mobility is known to depend on (1) the sign of the ionic charge, (2) the nature of the gas and particularly on impurities therein, (3) pressure, (4) temperature, (5) age of ion, and (6) can only within certain limits be regarded as independent of the magni-

tude of the electric field. Since the negative ion during part of its life may be simply an electron, its average mobility is in general higher than that of the positive ion. This is particularly evident with the lighter gases and gases at pressures below atmospheric, the ratio of the mobilities of the negative and positive ions continually increasing as the pressure is decreased. The mobility of the positive ions remains, over a considerable range, inversely proportional to the pressure.

Impurities are often found to affect the mobility of the negative ion to a remarkable degree. In gases such as argon, helium, and nitrogen, the negative ion apparently remains an unattached electron for a much greater proportion of time, if traces of water-vapour and electro-negative gases such as chlorine and oxygen are removed as completely as possible. Thus mobilities of the order of several thousand cm./sec. per V./cm. have been found in pure nitrogen with weak electric fields, although the normal value is less than 2 cm./sec. per V./cm. The ageing of positive ions, found in some cases, may also be attributed to the effect of impurities. In the first portion of its life, the ion may have a higher mobility than after a lapse of time of about 1 sec. In such cases it may be assumed that the ion has by that time encountered an impurity with which a particularly stable aggregate has been formed.

An important feature of ionization in mixed gases is that, for one sign of ion, normally only one mobility can be found, that is to say, the mobility of an ion depends in general only on the nature of the gaseous medium through which it moves and not on the particular gas in which it originated. Increase of temperature above normal produces negligible effect on the mobility calculated at constant density, but decrease of temperature causes a reduction in the mobilities of both positive and negative ions, the mobilities of the two kinds of ions becoming approximately the same.

Collision between ion pairs may result in effective neutralization, and the process of recombination of ions must usually be taken into account unless an electric field is present of sufficient strength to remove the ions from the gas, as they are formed, at such a speed as to allow them no appreciable opportunity for recombination. In the latter case, the maximum current possible apart from ionization by collision is obtained, that is, the saturation current. In small containing vessels loss of ions by diffusion may become important; the ions diffusing to the walls, where the charge is neutralized or lost. Ionic coefficients of diffusion for simple gases compare with those for diffusion of complex vapours into simple gases.

When ionization is produced by an α particle it is the direct result of a collision between the particle and a gas molecule. Ionization produced by X rays is only in small part immediately due to the expulsion of an electron from an atom on its absorption of a radiation quantum. Most of the ionization is produced by subsequent collisions of this electron with gas molecules. In general, ionization by collision is more readily brought about with electrons than with positive ions, since less potential is required to impart the necessary kinetic energy to an electron than to a positive ion, on account of the greater mobility of the former. At pressures from about 1 to 20 mm. of mercury, ionization, due to collision with gas molecules

of electrons derived from the gas itself, is easy to produce. The electric discharge current at constant potential then increases exponentially as the distance between the electrodes is increased, the exponential index, for a considerable range, being equal to the gas pressure multiplied by a function of the ratio of electric field to gas pressure.

A most important property possessed by an ion is that of facilitating the condensation of vapour. The Wilson expansion chamber, which utilizes this property to make visible the track of any ionizing particle passing through it, has from its inception been invaluable in the study of atomic phenomena. Recently[2], most remarkable photographs have been obtained with its aid of ionization phenomena due to cosmic rays. One shows, for example, the tracks of as many as 15 positive and 10 negative electrons ejected simultaneously from the nucleus of a lead atom as the result of absorption of a quantum of cosmic radiation. The energy value of the cosmic photon has in this way been shown to be of the order of $2\frac{1}{2}$ billion electron-volts.

§ 3. ABSORPTION OF X RAYS

The loss of intensity suffered by a beam of X rays in passing through a slab of material depends on (*a*) the initial distribution of intensity over the constituent wave-lengths and (*b*) the mass, atomic number and state of aggregation of the elements composing the slab. In the simple case of a homogeneous beam, approximately parallel, and having the one wave-length λ and initial intensity I_λ, the emergent intensity I after traversing a slab of uniform consistency and thickness d is given by $I = I_\lambda e^{-\mu_\lambda d}$, where μ_λ is a constant for the wave-length λ and a slab of definite uniform consistency. For a heterogeneous beam of limiting wave-lengths λ_1 and λ_2, the expression for the emergent intensity is then $\int_{\lambda_1}^{\lambda_2} I_\lambda e^{-\mu_\lambda d} d\lambda$. The constant μ is known as the *linear total absorption coefficient*.

λ, I_λ
I, d
μ_λ
μ

The process of X-ray absorption is fundamentally an atomic one. A quantum of energy $h\nu$ (where h is Planck's constant, $\nu = c/\lambda$, and c is the velocity of light) of the incident X-ray beam may interact with an atom in any one of the following ways: (i) (*a*) An electron having the kinetic energy $(h\nu - h\nu_K)$ may be expelled from the K level of the atom, and (*b*) a quantum $h\nu_K$ of characteristic radiation may leave the atom on the return of an electron to fill the K level. (ii) In addition to an electron having the kinetic energy $(h\nu - h\nu_K)$, other electrons, having kinetic energies such as $(h\nu_K - 2h\nu_L)$ may be expelled from the same atom[1]. In this case no K radiation appears. (iii) A quantum $h\nu$ of scattered X rays may appear. (iv) (*a*) A quantum $h\nu'$ of scattered X rays may appear, $h\nu'$ in this case being dependent in value on the angle of scattering (see p. 234); (*b*) simultaneously a recoil electron having the kinetic energy $(h\nu - h\nu')$ is ejected from the atom. The electron here affected must be regarded as so loosely bound prior to expulsion as to be practically free.

ν, c

The above processes combine to abstract energy from the incident X-ray beam and together result in the total absorption coefficient μ. The latter is often sub-

divided thus, $\mu=\tau+\sigma$, where τ measures the effect of the absorption processes (i) and (ii) and σ the scattering processes (iii) and (iv). Further, $\tau=\tau$ (true)$+\tau$ (fluorescent), where τ (true) includes the effect of (i) *a* and (ii) and τ (fluorescent) of (i) *b*. Similarly σ may be divided into σ (unmodified) due to (iii), and σ (modified) due to (iv); or alternatively, σ (true) due to (iii) and (iv) *a*, and σ_2 due to (iv) *b*, in which σ_2 really represents true absorption, just as does τ (true). τ σ

As ν increases, τ for any element decreases, until ν becomes equal to the excitation frequency of one of the main energy levels of the atom, when an immediate jump in the value of τ occurs, due to processes (i) and (ii) then becoming possible for the level in question. Discontinuities in the curve of absorption plotted against wave-length thus occur at wave-lengths corresponding to the K, L_1, L_2, L_3, M_1,... energy levels. Between any two successive discontinuities, a relation of the form $\tau=K.\lambda^n$ is found to hold. K, n

As soon as the work of Laue and Bragg had made it possible to employ homogeneous X rays, useful numerical values of the constants K and n for the commoner elements were found. An order of accuracy higher than about 1 per cent for these constants has, however, not been reached, mainly on account of difficulties connected with the measurement of X-ray intensity and with obtaining sufficiently thin and uniform sheets of pure material. Siegbahn[2] early found that n lay between 2·5 and 3, changing slightly for different elements. Richtmyer[3] and others have obtained results best represented by $n=3$, while Allen[4] and others have found n to be mostly 2·92 or thereabout. Universal formulae, such as $\tau/\rho=K.\lambda^3Z^4/W$ where ρ is density, Z atomic number and W atomic weight, have been proposed, in which K is a constant for any atom, so long as λ is short enough to excite the K radiation. Allen and Jönsson, however, have shown that these can only be roughly approximate. The best generalization is that given by Jönsson[5], who has shown that agreement with the bulk of experimental values for τ may be obtained with an accuracy probably not much different from that of the latter, by regarding $(\tau_e)_K$ as a continuous single-valued function of $(Z\lambda)$. In this expression $(\tau_e)_K$ is the electronic absorption coefficient τ_e, if the K absorption is being considered, or $\tau_e\times E_K/E_{L_1}$ in the case of the L absorption and $\tau_e\times E_K/E_{M_1}$ for the M absorption, while τ_e is connected with τ by the relation $\tau_e=\tau.A/\rho.L.Z$, in which L is Avogadro's number. ρ Z, W τ_e L

The multipliers E_K/E_{L_1} and E_K/E_{M_1} (i.e. $E_K/E_{L_1}\times E_{L_1}/E_{M_1}$) allow for the effect of the absorption discontinuities at the K and L edges. The extent of the K absorption jump, the total L jump and total M jump are better known[6] than the separate jumps at the L_2, L_3, M_2, M_3,..., edges, and it has been shown by Jönsson that the ratio of the two absorption coefficients at the K edge, for example, is equal to E_K/E_{L_1}, the ratio of the corresponding atomic energy levels.

On the short-wave-length side of the K and L absorption edges of a number of elements in crystalline compounds a fine structure has been observed. These maxima and minima of absorption occur within a wave-length range, equivalent to only a few hundred volts, from the absorption edge and were originally supposed to be simply due to the effect of chemical combination. A complete explanation of the

fine structure in crystals has, however, since been given by Kronig[7], who has shown that in this case it is a consequence of the existence of the energy field of the crystal lattice.

In any application of the above formulae, it should not be overlooked that they apply only when a narrow, nearly parallel, X-ray beam is used, while for wide angle beams, especially with short wave-lengths, much scattered radiation may be included in the emergent cone of rays, considerably enhancing its intensity over values calculated as above.

Theories of X-ray absorption have been worked out by De Broglie[8], Thomson[9], Compton[10], Kramers[11] and Wentzel[12]. Wentzel's theory is based on Schrödinger's wave-mechanics. Unfortunately no theory yet exists capable of yielding sufficiently accurate numerical values of μ to be of much practical use. Reference must be made to separate experimental values and empirical formulae.

§4. SCATTERING OF X RAYS

Radiation proceeding from an atom, as a result of interaction between the latter and a beam of homogeneous radiation incident upon it, is in general of composite character. Part of it consists of radiation having wave-lengths characteristic of the element concerned and belonging to the K, L, M,... series of spectral emission lines. The remainder consists of (*a*) radiation of the same wave-length as the incident radiation, together with (*b*) radiation composed of slightly longer wave-lengths. Both (*a*) and (*b*) are now regarded as scattered radiation, (*a*) being referred to as "unmodified" or "coherent" and (*b*) as "modified" or "incoherent." The main properties of (*b*) were discovered by Compton[1], although it had been shown some years earlier by Sadler and Mesham[2] that the average wave-length of scattered radiation was greater than that of the primary.

Z The term "incoherent" has been applied to the modified radiation, since it is found that in order to arrive at the total intensity of such radiation emitted by an atom containing a number Z of extra-nuclear electrons it is only necessary to add the separate intensities of this kind of radiation emitted by each of these electrons. To obtain the total unmodified intensity, however, the wave amplitudes of the unmodified radiation due to each of these electrons must be added with due regard to phase relationship, and the square of the resultant amplitude furnishes a measure of the total intensity, so that this radiation behaves as coherent.

Compton showed that his results could be accounted for in a natural manner on the quantum theory of radiation by assuming that the radiation quantum or photon
h of energy $h\nu$ also possessed momentum of amount $h\nu/c$, where h is Planck's con-
ν, c stant, ν the frequency of the incident radiation, and c the velocity of light. The interaction between the photon $h\nu$ and the atom, resulting in a scattered photon of
ν', ϕ smaller energy $h\nu'$ at an angle ϕ with the direction of the incident photon and causing the ejection of a "recoil" electron, could then be regarded simply as a collision between the photon $h\nu$ and this electron. Applying the laws of conservation of

energy and momentum, and using the relativistic expressions for the momentum and kinetic energy of the electron, Compton obtained the equation

$$\lambda' - \lambda = \frac{h}{m.c}(1 - \cos\phi),$$

where λ and λ' are the wave-lengths of the incident and scattered radiation respectively and m is the electronic mass. λ, λ' m

After the discovery of the Raman effect, a corresponding "anti-Stokes" radiation was looked for in the X-ray region and some results were obtained[3] which were thought to show that it was possible for an X-ray photon to acquire energy in the process of scattering, instead of losing it. Other investigations[4], however, have shown that there appears to be no evidence of this, nor, in fact, of anything corresponding to any part of the Raman effect.

Absorption measurements of scattered radiation will in general show the existence of longer wave-lengths than those calculated from Compton's equation given above. This result arises from the presence of radiation that has more than once been scattered, and with elements of low atomic number and incident radiation of short wave-length, the effect of double scattering may be very appreciable.

Theoretical expressions for the intensities of both the modified and unmodified components of radiation arising from atomic scattering have been built up by several authors and have been found[5] to agree well with experiment. One due to Compton[6] is:

$$I_\phi = I_e \left\{F^2 + \left(Z - \frac{F^2}{Z}\right)(1 + \alpha \text{ vers } \phi)^{-3}\right\},$$

where

$$I_e = \frac{I_0 e^4}{2m^2c^4}(1 + \cos^2\phi), \quad F = \int_0^\infty U(r)\frac{\sin kr}{kr}.dr, \quad k = \frac{4\pi}{\lambda}.\sin\frac{\phi}{2}, \quad \alpha = \frac{h\nu}{mc^2}, \qquad I_e, F, k, \alpha$$

$U(r)$ is the radial charge density of the atom measured in electrons per unit distance, $U(r)$
e is the electronic charge, I_0 is the intensity of the incident radiation, assumed unpolarized, and I_ϕ is the total intensity of the radiation scattered by one atom per unit solid angle in a direction making an angle ϕ with the incident radiation. The coherent part of the scattered radiation is represented by the term I/F^2. The expression for I_e, which represents the scattering by a single electron for relatively long X-ray wave-lengths, i.e. for unmodified scattering, was deduced by Sir J. J. Thomson[7]. The factor $I_e(1 + \alpha \text{ vers } \phi)$ represents the scattering by a free electron when only modified scattering takes place, and has been obtained by Dirac[8] on the basis of quantum mechanics. A linear coefficient of scattering may be defined as the scattered intensity per unit solid angle per unit length of scatterer traversed by a parallel incident beam of unit intensity. The linear coefficients s_1 and s_2 of coherent and incoherent scattering, respectively, are then given by

e, I_0 I_ϕ I_e

$$s_1 = KF^2 \quad \text{and} \quad s_2 = K(Z - F^2/Z)(1 + \alpha \text{ vers } \phi)^{-3}, \qquad K$$

where $K = \frac{Le^4.\rho}{2m^2c^4.W}(1 + \cos^2\phi)$, L is Avogadro's number, W the atomic weight and ρ the density of the scattering element. L, W ρ

In many cases in practice, the phenomena of scattering are still further compli-

cated by interference between radiation scattered by the different atoms of the scattering substance. Thus even with a gaseous scatterer, if radiation of sufficiently long wave-length is used for the scattered radiation to be of conveniently measurable amount, it is found[9] that the grouping of atoms into molecules gives rise to diffraction maxima and minima, which greatly alter the form of the distribution of intensity with scattering angle from that due to purely atomic scattering. Similar effects are observed with liquid scatterers. Scattering from any of the many solid materials consisting of aggregates of small crystals is affected to a marked extent by crystalline interference phenomena (see p. 239) unless the wave-length of the incident radiation is sufficiently short in relation to the angle of scattering considered[5].

§ 5. MEASUREMENT OF INTENSITY, QUALITY, ETC.

Intensity. Since a beam of X rays consists of a train of electromagnetic waves, it would appear logical to define X-ray intensity in terms of energy flux or energy density in a manner analogous to that adopted in the case of other systems of waves. Although such a definition would have the advantage of being theoretically sound, there is as yet no method by which absolute measurements may be made conveniently and with sufficient accuracy. Further, it would appear at present impossible (even if it were desirable) to adopt a system of units based on a standard source analogous to the standard candle in light.

In the absence of satisfactory absolute methods of measuring X-ray intensities, recourse is had to indirect measurements, advantage being taken of the chemical or ionizing effects of the rays. Among chemical methods the change in colour of barium platino-cyanide has been extensively used in medical practice. Crystals of barium platino-cyanide, normally green in colour, change to a brown tint when exposed to X rays, and this colour change, expressed in arbitrary units, has been related to the biological effects of the radiation[1]. Similar measurements may be made by observations on the degree of blackening produced by the X rays in a photographic film. Measurements by these methods are not susceptible of a very high accuracy, although Bouwers[2] has shown that by taking suitable precautions reasonably good results may be obtained photographically. A more satisfactory basis of measurement is found in the ionization of gases, and ionization methods have now been developed to a considerable accuracy.

All these indirect methods appear to give a measure of the amount of radiation absorbed by the measuring pastille, film or ionization chamber, so that unless corrections are applied they determine not the energy or intensity of the beam but a quantity resting only partly on the intensity and partly on the quality. In the case of beams of known composition such measurements can be converted into energy-measurements, but when we are dealing with heterogeneous beams of unknown composition this conversion is impracticable.

The ionization of air by X rays is often directly associated with the dose of therapeutic practice, since air absorption is commonly accepted as affording a physical measure of biological effect, and a working standard of X-ray quantity or dose on this basis was adopted by the Second International Congress of Radiology

(1928). The unit so adopted is the röntgen (*r*) defined as "that quantity of X radiation which, when the secondary electrons are fully utilized and the wall effect of the chamber is avoided, produces in 1 cm³ of atmospheric air at 0° C. and 76 cm. mercury pressure, such a degree of conductivity that 1 e.s.u. of charge is measured at saturation current." (*r*)

No special design of ionization chamber is suggested in the definition and, in the various attempts which have been made to realize the unit, several types of chambers have been developed. These may be divided into two classes, namely, *free-air*

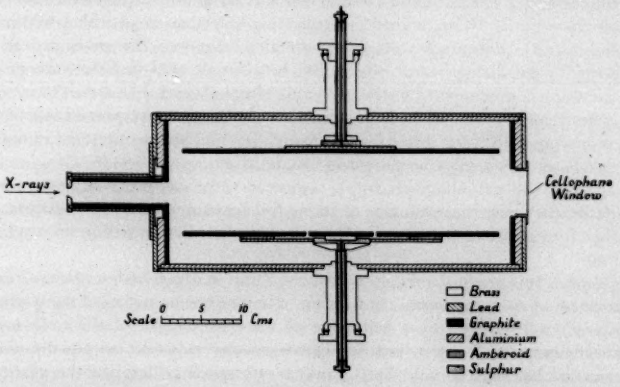


Figure 1. Standard free-air ionization chamber at the National Physical Laboratory.

chambers[3] in which the volume of air ionized is not bounded by walls, and *air-wall* chambers[4] in which the volume of air ionized is bounded by conducting walls having as nearly as possible the same properties as air as regards X-ray absorption and scattering. It would appear that for X rays excited at voltages up to about 200 kV. peak, free-air chambers provide reliable and accurate primary standards, while air-wall chambers, properly calibrated, furnish satisfactory and convenient working standards.

A typical free-air ionization chamber is shown diagrammatically in figure 1. The two electrodes are parallel plates of graphite supported on aluminium plates mounted symmetrically in an aluminium case. Graphite is used to minimize wall effects. The interior of the case is also coated internally with graphite and externally with lead to shield the interior from both direct and scattered radiation. One of the electrodes is raised to a potential of several hundred volts. The other electrode consists of a collector surrounded by a co-planar earthed guard-plate.

The volume of air ionized may be deduced from the geometry of the X-ray beam, or by photographic means, and the quantity of electricity collected on the guarded electrode may be measured, for example, by the use of a suitable electro-

meter. These measurements serve to permit the quantity of X rays to be measured in röntgens. Certain practical details regarding diaphragms etc. must be observed in work of the most accurate type.

In 1931 an intercomparison of the primary standard dosemeters of the National Physical Laboratory, the National Bureau of Standards, and the Physikalisch-Technische Reichsanstalt, showed an agreement within 0·5 per cent[5].

Air-wall chambers consist essentially of an envelope of material having an atomic number approximately the same as that of air (e.g. celluloid rendered conducting by a coating of a mixture of graphite and aluminium oxide) which serves to define the volume of air ionized. A collecting electrode is provided within the envelope, and a saturation voltage is maintained between the envelope and the collector. A measuring device, which may be a simple gold-leaf electroscope or a valve amplifier, is connected to the collecting electrode and thus serves to measure the quantity of X rays absorbed. Instruments of this type have proved satisfactory for measuring hard X rays, but for soft rays difficulties are encountered in making the envelope sufficiently transparent to the radiation.

Although air-wall chambers for X rays need to be calibrated against primary standards, they have the advantage of taking full account of scattered radiation, and are therefore of use in determining the intensity-distribution within an irradiated medium.

Quality. It is proving difficult to establish a unit in which to describe the quality or hardness of an X-ray beam. In the case of homogeneous radiation the quality is sufficiently well defined by a statement of the wave-length of the radiation in angstrom units, or X units, but for heterogeneous radiation no equally simple specification has been found. Early workers attempted to describe the quality of the radiation by a statement of the exciting voltage on the X-ray tube (for example, the alternative spark gap) but it has now been established that such a unit is inadequate[6]. Later it was suggested that quality might be defined in terms of the thickness of some suitable absorber which reduces the intensity to half its original value (the so-called "half-value layer") or by the absorption curve of the radiation in some suitable material. It would appear at present that the only reasonably satisfactory method of specifying quality is by a statement of generating conditions including the peak or r.m.s. value and type of the exciting voltage together with some data concerning the absorption of the radiation.

X-ray protection. After a series of casualties occurring to X-ray and radium workers a committee was formed in 1921 to draw up recommendations concerning the safety and well-being of such workers. This committee, known as the British X-ray and Radium Protection Committee and consisting of representatives of interested bodies, both physical and medical, under the chairmanship of Sir Humphry Rolleston, issued its first recommendations in 1921, and successive revised reports in 1923, 1927, and 1934.* Since its inception the committee has worked in

* Copies of the International Recommendations for X-ray and Radium Protection and of the Recommendations of the British X-ray and Radium Protection Committee may be had on application to the Director, National Physical Laboratory, Teddington, Middlesex.

close co-operation with the National Physical Laboratory, which has carried out a series of investigations on the various problems associated with X-ray and radium protection[7].

A set of International Recommendations, issued by the International X-ray and Radium Protection Committee, were adopted by the Second International Congress of Radiology in 1928 and revised at the International Congresses of 1931 and 1934. These and somewhat similar recommendations published by the Health Organization of the League of Nations[8] in 1931 are largely based on the British proposals. The principles underlying X-ray and radium protection are: (1) the maintaining of the general health of the worker by the provision of light and well-ventilated rooms, adequate holidays, etc.; (2) the reduction of the exposure to radiation by the use of suitable protective materials*; and (3) the reduction of the risks of electrical fatalities by the proper design of high tension apparatus.

These various recommendations, which have been followed in most present-day hospital and industrial X-ray departments, have done much to reduce the risks attending X-ray and radium work, and it would appear that X-ray dermatitis and similar diseases are practically unknown among workers who have continually observed the recommendations.

§ 6. X-RAY SPECTRA

Introduction. The discovery of X rays by Röntgen was followed by many experimental researches into the properties and the nature of the rays, but a very considerable time was to elapse before it was finally and definitely established that these rays were similar to light waves and distinguished from them merely by their shorter wave-length. Many attempts were made to detect refraction, reflection and diffraction but in only one case was a positive result obtained. In 1902 Haga and Wind observed evidence of diffraction through a narrow slit but the effect was so extremely small that the experiment was not considered as definitely conclusive. Barkla's discovery of the polarization of scattered X rays lent further support to the view that the rays were electromagnetic waves but it was not until 1912 that their nature was settled beyond question.

In that year Laue conceived the brilliant idea of using a crystal as a diffraction grating. The previous experimental work had shown that, if X rays were similar to light waves, their wave-length must be extremely short and probably of the order of 10^{-8} or 10^{-9} cm. This distance is of the same order as the distance separating adjacent atoms in a crystal lattice and Laue realized that such a lattice should serve as a three-dimensional diffraction grating for X rays. He investigated the mathematics of the problem and showed that the passage of a fine pencil of X rays through a crystal should give rise to a diffraction pattern. With Friedrich and Knipping, Laue carried out the experiment and his mathematical predictions were precisely confirmed.

This work of Laue's was followed up by W. H. Bragg and W. L. Bragg, who con-

* From evidence at present available it would appear that under satisfactory working conditions a person in normal health can tolerate exposure to X rays to an extent of about 0·2 röntgen per day.

sidered the problem of the interaction of the rays with the regularly arranged atoms of the crystal from a different point of view. They showed that Laue's results could be considered as arising from the reflection of the X rays by the various planes or sets of planes of the crystal. They proved that the reflected beam was in the same plane as the incident beam and the normal to the reflecting plane, and that the angle of incidence was equal to the angle of reflection but that reflection only occurred when a further condition was satisfied. This condition, now known as the Bragg law, is expressed by the equation

$$n\lambda = 2d \sin \theta,$$

λ, d, θ, n where λ is the wave-length, d the distance separating successive identical planes of atoms in the crystal, θ the glancing angle which the incident or reflected beam makes with the plane and n an integer giving the order of the reflection. Only when this condition is satisfied are the waves reflected from successive planes of atoms in phase with each other and capable of giving rise to a reflected beam.

W. H. and W. L. Bragg investigated certain simple crystal structures such as, for example, rock salt. In this they found that the X-ray reflections were consistent with a cubic structure in which the chlorine and sodium atoms were placed at alternate cube corners. From a knowledge of the density and of Avogadro's constant they deduced that the distance separating successive identical planes parallel to the cube face was $2{\cdot}814 \times 10^{-8}$ cm. The d of the Bragg equation was thus known and the way was opened for accurate wave-length determinations in the X-ray region.

This work of Laue and of W. H. and W. L. Bragg has developed in two main directions. In the first place it has led to our very detailed and accurate knowledge of the X-ray spectra of the elements, a knowledge which has contributed vastly to the elucidation of many of the problems of atomic structure. In a different but equally important direction it has led to the development of the subject of X-ray crystal analysis and in this way has had far-reaching effects in both the pure and the applied sciences.

The X-ray spectra of the elements. It is to Barkla's experiments that we owe our first knowledge that each element has associated with it X rays of particular and characteristic quality. Barkla showed that, when X rays fell on an element, among the rays scattered by the material there existed a preponderance of rays with a definite penetrating power. Absorption experiments of these rays in, say, aluminium proved their homogeneous nature, and in a series of classical experiments Barkla showed that to each element there were to be ascribed two characteristic radiations, one of which was much more penetrating than the other. This penetrating power increased with atomic weight and was independent of the state of chemical combination. To these he gave the names of *K* and *L* radiations. Until the work of Laue and W. H. and W. L. Bragg the only way in which these radiations could be described was in terms of their absorption coefficients, but the discovery of interference methods of measuring X-ray wave-lengths enabled Barkla's work to be extended and expressed in terms of wave-length.

The first series of measurements of X-ray wave-lengths came in 1913 when Moseley published his first paper on "The high frequency spectra of the elements."

In these researches he used a crystal face of potassium ferrocyanide with a grating space for which $d = 4{\cdot}227 \times 10^{-8}$ cm. Moseley's experiments on the lighter elements (calcium to zinc) showed that the K radiation was not monochromatic but consisted of two principal lines, a strong α line and a weaker β line. A later paper showed that the L series was even more complex. These papers of Moseley's were outstanding not only in giving the first proper measurements of X-ray wave-lengths but also in indicating the nature of the essential laws governing X-ray emission spectra. He showed for example that the variations of wave-length from one atom to another depended upon the atomic number and not on the atomic weight, and that the square root of the frequency of any line was a linear function of the atomic number. For example, for the α line he found that

$$\frac{\nu}{R} = \frac{3}{4}(Z-1)^2 = (Z-1)^2\left(\frac{1}{1^2} - \frac{1}{2^2}\right),$$

where ν was the frequency of the line, Z the atomic number of the emitting element, and R the Rydberg constant of optical spectrometry. Similar relations were shown to hold for other lines of the K and L series. These results, coming as they did just when Bohr had put forward his theory of the hydrogen spectrum, afforded valuable support to that theory and showed that the X-ray emission lines could be considered as the result of the passage of an electron from one quantum state to another. At the same time Moseley was able to assign to the various elements their appropriate atomic numbers and to show what elements remained to be discovered. In the subsequent discovery and identification of certain of these elements, X-ray spectroscopy has fittingly played an important part.

ν, Z
R

Since this pioneer work of Moseley, our knowledge of the details of X-ray spectra has been vastly increased. To-day, thanks largely to the work of Siegbahn and his school, we have a very complete and accurate knowledge of the X-ray spectra of the elements and, arising from these measurements, a very detailed knowledge of the distribution of electrons and energy-levels in the various atoms.

The X-ray emission spectra. The X-ray emission spectrum of an element consists of several well-defined groups of lines to which, following Barkla's lead, the letters $K, L, M, \ldots$ have been ascribed. As typical, the X-ray spectrum of tungsten consists of a K series with lines varying in wave-length from 0·178 to 0·213 A., an L series covering the range 1·025 to 1·675 A. and an M series from 5·162 to 8·975 A. Certain lines in the N series with a wave-length of 56 A. have also been observed. The K spectrum has now been measured from lithium ($Z = 3$) to uranium ($Z = 92$), the L series from sodium ($Z = 11$) and the M series from selenium ($Z = 34$) upwards. This means that practically the whole available range has been investigated.

The K series consists in general of a single group of five lines, although in the case of the lighter elements this number is considerably increased. The principal lines include a strong close doublet α_1, α_2, a second doublet β_1, β_3 and a weaker β_2 line. The L series shows three groups of lines. In the first group the chief lines are (using Siegbahn's notation) β_3, β_4, γ_2, γ_3 and γ_4; the second group contains η, β_1, γ_5, γ_1, γ_8 and γ_6, while the third group includes l, α_2, α_1, β_6, β_2, β_7 and β_5. These

three groups require different potentials to excite them. Thus, for platinum the whole L spectrum is excited by a potential of 12,000 V., the second and third groups only by a potential of 11,600 V., while if the voltage is reduced to 10,200 only the third group appears. The M series is even more complicated and contains five groups of lines. The measurements in this series are not so complete as those in the shorter wave-length series since here, owing to the long-wave lengths involved, the experimental difficulties are much greater. In the $N, O, \ldots$ series only isolated observations have been made for the heavier elements.

The X-ray absorption spectra. The X-ray absorption spectra of the elements are not, in general, characterized by lines such as are found in the optical region but by the existence of absorption edges sharply defined on the long wave-length side. Corresponding to the $K, L, M, \ldots$ emission spectra there exist $K, L, M, \ldots$ absorption edges. The K absorption edge is single and occurs at a wave-length just shorter than the shortest of the K emission lines. There are three L absorption edges $L_{\mathrm{I}}, L_{\mathrm{II}}, L_{\mathrm{III}}$, corresponding to the three groups of lines in the L emission spectrum. In the same way there are found to be five M edges. The N and O absorption edges have not been actually measured but we have indirect knowledge that the N absorption spectrum, for example, must consist of seven edges. When the frequencies corresponding to these edges are measured it is found that the variation with atomic number is similar to that found for the emission lines, so that the Moseley curves, where $\sqrt{(\nu/R)}$ is plotted against the atomic number Z, are approximately straight lines. Accurate measurements have proved that the square root of the frequency is not a purely linear function of the atomic number but shows certain discontinuities. It will be shown later that these discontinuities have an important bearing on the problem of the distribution of the electrons in the various atoms.

The interpretation of X-ray spectra. It will be clear from this brief description of the nature of X-ray emission and absorption spectra that there is a marked difference between the spectrum of an element in the X-ray region and its optical spectrum. For example, the optical spectra of the elements such as the alkalies which occupy corresponding positions in the periodic table bear a close resemblance to one another and are sharply differentiated from those of the alkali earths. On the other hand, in the X-ray region the spectra of all the elements are similar to one another and differ only in a progressive increase in frequency with increase in atomic number.

In considering the interpretation of X-ray spectra the following considerations are of importance. In the first place the X-ray spectra are practically independent of the state of chemical combination so that it must be the inner electrons which play an important part. This is also shown by the high frequencies associated with the X-ray lines, since such high frequencies imply correspondingly high energies and suggest that the innermost and most tightly bound electrons are participating in the production of such spectra. Also the X-ray spectrum consists of several well-defined $K, L, M, \ldots$ groups of lines. In any element the frequencies of these groups are widely separated so that we are forced to the conclusion that the electrons in an atom can be sorted out into groups with correspondingly widely separated energies. The similarity of the X-ray spectrum of one atom to that of its neighbours shows that, so

far as the inner groups of electrons go, the structure of one atom must closely resemble that of its neighbours.

The first step towards the emission of a line in the X-ray spectrum is the ejection of an electron from one or other of the inner groups. This may result from an electron collision or may occur as a consequence of the absorption of energy of a primary X-ray. The electron ejected may belong to the K group or to one of the lower energy L, M, ... groups. The second step is the filling up of the gap produced by the ejection of this electron. This can be accomplished in several ways. For example, if the gap occurs in the K level it can be filled by the passage of an electron from the L group or by one from the M, N or one of the outer groups. The associated radiation will have a different energy according to the source of the electron. Thus if the line occurs as a result of the passage of an electron from an L group to the K group, it will have a lower frequency than if the transition involved was from an M group, since the energy difference between the M group and the K group is greater than that between the L and K groups. The intensities of the lines will depend on the relative probabilities of the various possible transitions.

This simple interpretation provides a very exact explanation of the observed frequencies of the great majority of X-ray lines. A correspondingly simple interpretation accounts for the absorption edges. If ν_K is the frequency of say the K absorption edge, then $h\nu_K$ is the energy necessary to remove the electron from the K level to the outside of the atom. The absence of a true absorption line spectrum is explained by the fact that the only possible transitions for an electron removed from, say, the K group are either to the outside of the atom or at any rate to one of the outermost levels which may not be fully occupied. In general the inner levels will be fully occupied by electrons and therefore offer no resting place for the ejected electrons, so that absorption lines with frequencies corresponding to the emission lines will not be observed.

ν_K

If $\nu_K, \nu_{L_I}, \nu_{L_{II}}, \nu_{L_{III}}, \ldots$ are the frequencies associated with the $K, L_I, L_{II}, L_{III}, \ldots$ absorption edges, the corresponding energies $h\nu_K, h\nu_{L_I}, h\nu_{L_{II}}, h\nu_{L_{III}}, \ldots$ measure the energies necessary to remove an electron from the $K_I, L_I, L_{II}, L_{III}, \ldots$ levels to the outside of the atom. If $W_K, W_{L_I}, W_{L_{II}}, W_{L_{III}}, \ldots$ represent the energies of the atom when an electron has been removed from the $K, L_I, L_{II}, L_{III}, \ldots$ levels it follows that when the atom emits a $K\alpha$ line, the energy before emission is W_K and after emission W_L, so that in accordance with the Bohr relation

W_K

W_L

$$h\nu_{K\alpha} = W_K - W_L.$$

Similarly for a $K\beta$ line, when the electron comes direct from an M level to fill the gap in the K level,

$$h\nu_{K\beta} = W_K - W_M.$$

For an $L\alpha$ line, the gap in the L level is filled by an electron from the M level, so that

$$h\nu_{L\alpha} = W_L - W_M.$$

From these equations the following relations must connect the frequencies of the K and L emission lines:

$$\nu_{K\beta} = \nu_{L\alpha} + \nu_{K\alpha},$$

a relation which was very early found to be satisfied. This relation can be seen otherwise. An electron is removed from the K level, its place can be filled in several ways. Either an electron comes straight to the K level from the M level and the $K\beta$ line is emitted; or the gap in the K level may be filled by an electron from the L level with the emission of the $K\alpha$ line, and the gap thus made in the L level is filled from the M level and the $L\alpha$ line is emitted. A consideration of the energetics of the two processes leads to the equation given above.

When we consider the complexity of the various levels we see that they are capable of giving many more lines than are actually observed. Thus, with one K, three L, five M and seven N levels we should expect fifteen lines in the K series without considering the contributions of the outer N and O levels. For the heavier elements only five lines are observed and a numerical comparison shows that these arise from the following transitions:

$$K\alpha_1 = K - L_{III},$$
$$K\alpha_2 = K - L_{II},$$
$$K\beta_1 = K - M_{III},$$
$$K\beta_3 = K - M_{II},$$
$$K\beta_2 = K - N_{II}, N_{III}.$$

No lines occur, for example, corresponding to the transitions $K-L_I$, $K-M_I$, ... To obtain an explanation of the occurrence of certain lines and the non-occurrence of others we must consider the quantum relations corresponding to the various transitions. On modern theory each atomic state is characterized by four quantum numbers n, l, j and s. n is the principal quantum number, l the azimuthal quantum number, j the internal quantum number which gives the resultant angular momentum of the electron and its spin; while the number s, which for such transitions as are involved in X-ray spectra can have the values $\pm\frac{1}{2}$, is associated with the spin of the electron. The numbers l, j and s are connected by the relation $j = l \pm s$, while the quantum number l must always be less than n. The K, L, M, N, ... levels are associated with successive values 1, 2, 3, 4, ... of n. For the K level we have $l=0$, while j has the value $\frac{1}{2}$. For the L level, where $n=2$, l may have the values 0 or 1. To the zero value of l corresponds the value $\frac{1}{2}$ of j, while for the case when $l=1$, j may be either $\frac{1}{2}$ or $\frac{3}{2}$. We thus find three L levels with n, l, j values of $20\frac{1}{2}$, $21\frac{1}{2}$ and $21\frac{3}{2}$. Similarly for the M levels the values of n, l, j are $30\frac{1}{2}$, $31\frac{1}{2}$, $31\frac{3}{2}$, $32\frac{3}{2}$, $32\frac{5}{2}$, giving five possible levels as are actually observed. The emission lines correspond to transitions between these levels but are subject to certain selection rules. Modern wave-mechanical theory shows that the transitions which are to be expected are those in which (1) n alters, (2) l changes by ± 1, and (3) j either remains unaltered or alters by ± 1.

n, l
j
s

These energy-levels together with this selection principle provide a satisfactory explanation of the vast majority of the observed emission lines. In figure 2 is given a typical energy-level diagram for the uranium atom ($Z=92$). The arrow-heads denote the transitions which give rise to the various emission lines.

Doublets in the X-ray spectrum. Certain interesting results follow immediately from such atomic models. For example, the $K\alpha_1$, $K\alpha_2$ doublet arises from the transitions $K-L_{III}$ and $K-L_{II}$. The energy difference, or the frequency difference for such a doublet should be identical with those of the two absorption edges L_{II} and L_{III}. A study of the diagram will show that certain pairs of lines in the L series should also show this same frequency difference. Such pairs are (η, l), (β_1, α_2), (γ_5, β_6), (γ_1, β_{15}) and (γ_6, β_5), since for each of these pairs the frequency difference is

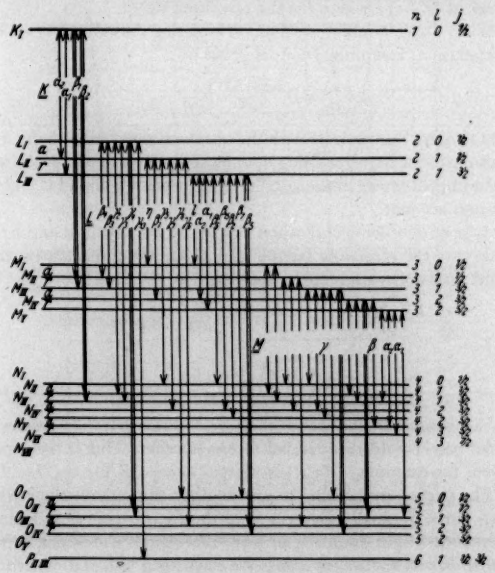


Figure 2. Energy-levels for the uranium atom

that of the L_{II}, L_{III} levels. This is completely confirmed by actual measurements of the frequencies. Doublets of this kind which correspond to transitions either from or to two levels with the same value of l but different values of j are known as spin doublets.

In accordance with wave mechanics the formula giving the energy W in the quantum state n, l, j, s is

$$\frac{W}{Rh}=\frac{(Z-s_{nl})^2}{n^2}+\alpha^2\frac{(Z-d_{nlj})^4}{n^3}\left(\frac{1}{l+\frac{1}{2}}-\frac{3}{4n}\right)-\alpha^2\frac{(Z-d'_{nlj})^4}{n^3}\left(\frac{j(j+1)-l(l+1)-s(s+1)}{2l(l+\frac{1}{2})(l+1)}\right).$$

Here Z is the atomic number; s, d, d' are screening constants, R is the Rydberg constant, while α, which has the value $2\pi e^2/ch$, arises from the relativity correction. Z, s, d, d', R α

The screening constants arise from the following considerations. The action of the nuclear charge Z on an electron in, say, an L level is affected by the inner K electrons which, to a first approximation, behave as if they were at the nucleus and reduce the effective nuclear charge to $Z-2$. It is also affected by the other L electrons and by the outer M, N,... electrons. The outer electrons contribute little to the effect since they behave approximately as an outer conducting spherical shell. The resultant of all these effects is that the electron is acted on by a central force corresponding to a
s nuclear charge of $(Z-s)\,e$ where s is the screening factor.

If we consider a spin doublet, we find, on inserting the values of $j=l\pm\frac{1}{2}$, that the frequency separation, assuming $d=d'$, is given by

$$\Delta\left(\frac{\nu}{R}\right)=\alpha^2\frac{(Z-d)^4}{n^3}\left(\frac{1}{l\,(l+1)}\right).$$

This is found to agree numerically with the observed separation in the whole X-ray region and also with the corresponding doublet separation in the optical spectrum of hydrogen. An improvement in the numerical agreement is found if higher terms in α^2 are taken into account.

Another type of doublet occurs when l varies by a single unit and n remains the same. In this case the screening factors s_{nl}, s_{nl}' are the important factors and for such a doublet, known as a screening doublet,

$$\frac{\nu}{R}=\frac{(Z-s_{nl})^2}{n^2}\quad\text{and}\quad\frac{\nu'}{R}=\frac{(Z-s_{nl}')^2}{n^2}\quad\text{approximately,}$$

so that

$$\Delta\sqrt{\left(\frac{\nu}{R}\right)}=\frac{s_{nl}-s_{nl}'}{n}.$$

Since s_{nl}, s_{nl}' are practically independent of Z it follows that the Moseley curves for such a doublet pair should run parallel to one another. This is confirmed experimentally when, for example, $\sqrt{(\nu/R)}$ is plotted against Z for the L_{I}, L_{II} absorption edges. The curves are parallel to one another throughout their whole length while the L_{III} curve gradually diverges from the L_{I}, L_{II} pair as is to be expected since L_{II}, L_{III} form a spin doublet pair whose frequency difference depends on Z.

The electron distribution in the atom. We have seen that the experimental data obtained from X-ray emission and absorption spectra indicate the existence of various energy-levels in the atom, one K level, three L levels, five M levels, seven N levels, etc. A more detailed study of the numerical data results in an accurate picture of the way in which such levels are built up as we progress from atom to atom in the periodic table. It has already been pointed out that the Moseley curves in which $\sqrt{\nu/R}$ is plotted against Z are approximately straight lines. Accurate measurements, however, show that their lines show certain discontinuities and these discontinuities are intimately concerned with the formation of successive levels. This feature of the Moseley curves is emphasized if instead of using $\sqrt{\nu/R}$ as
a, b ordinate we use $\sqrt{\nu/R}-(aZ+b)$, where a and b are so chosen that this quantity is small. Such curves consist of a series of approximately straight lines, the various straight portions corresponding to the filling up with electrons of some particular

level. A typical curve is reproduced in figure 3. An examination of the L_{II}, L_{III} curves shows a continuous line up to $Z=10$ when the L levels are completed, other straight portions from 13 to 18 while the M_{II}, M_{III} levels are being filled, from 21 to 29 (the iron group) corresponding to the M_{IV}, M_V levels and so on. By a study of curves such as this the number of electrons in each level can be determined and also

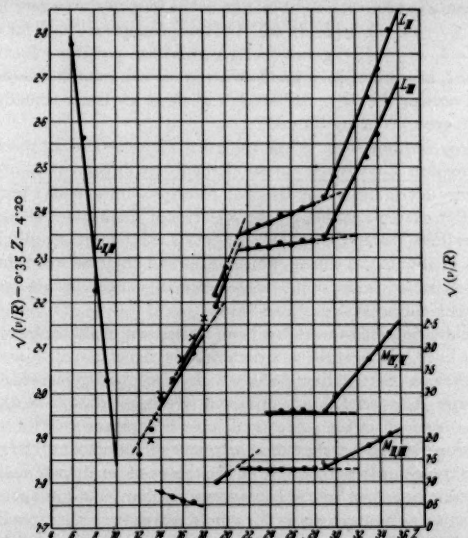


Figure 3. Moseley curves relating $\sqrt{(\nu/R)}-(aZ+b)$ to the atomic number Z.

the order in which the levels are filled up as Z is increased. The final result for the number of electrons in the various levels of the uranium atoms is for example:

K	L_I ($L_{II}L_{III}$)	M_I ($M_{II}M_{III}$) ($M_{IV}M_V$)	N_I ($N_{II}N_{III}$) ($N_{IV}N_V$) ($N_{VI}N_{VII}$)
2	2 6	2 6 10	2 6 10 14

O_I ($O_{II}O_{III}$) ($O_{IV}O_V$)	P_I ($P_{II}P_{III}$) ($P_{IV}P_V$)	Q_I
2 6 10	2 6 4	2

The distribution of electrons obtained in this way throws interesting light on many problems connected with the periodic table. It is found for example that the O_I level is occupied while the N_{IV}, N_V, N_{VI}, N_{VII} levels are still empty. The N_{IV}, N_V levels are next filled followed by the O_{II}, O_{III} and P_I levels, and it is not until cerium is reached ($Z=58$) that electrons begin to enter the N_{VI}, N_{VII} levels.

Between $Z=58$ and $Z=71$ these levels are being filled up, and we have an explanation of the occurrence of the rare earth group in which the elements are characterized by their chemical similarity.

The distribution obtained in this way from a study of the X-ray data is in excellent agreement with that which follows from Pauli's principle that not more than one electron can exist for each combination of the four quantum numbers n, l, j and m where $m = \pm j, \pm(j-1), \pm(j-2)$, etc. On this principle we have for the M_{IV} level $n=3$, $l=2$, $j=\frac{3}{2}$, $m=\pm\frac{3}{2}$, $\pm\frac{1}{2}$ and four electrons are possible. For the M_V level $n=3$, $l=2$, $j=\frac{5}{2}$, $m=\pm\frac{5}{2}$, $\pm\frac{3}{2}$, $\pm\frac{1}{2}$ and six electrons can exist in this level. Thus ten electrons can occupy the M_{IV}, M_V levels and these levels are actually being filled while the iron group is being formed.

Other X-ray emission lines. While the great majority and all the more intense lines of the emission X-ray spectra of the elements are directly accounted for by this system of energy-levels and by the selection principle, other weaker lines fall outside these rules. For example, although the K series of the heavier elements contains only about five lines, that of the lighter elements is much more complicated. Thus in the K spectra of aluminium silicon, phosphorus and chlorine some eleven lines are found, and even in the spectra of the heavier elements weak lines are observed which do not fit in with the scheme outlined above.

A considerable volume of work has been carried out in attempts to elucidate the origin of these lines. Such work has shown that the selection rules quoted above are not always sufficient. Certain lines occur which correspond to forbidden transitions. Thus weak lines $K\beta_4$ and $K\beta_5$ are observed which correspond to the transitions $K-(N_{IV}, N_V)$ and $K-(M_{IV}, M_V)$ and have been measured for the elements rubidium to tin. For these transitions the quantum number l undergoes a double jump. Such a transition is permissible on the theory of quadripole radiation, which permits of double jumps of both the quantum number l and j as against the more restricted selection rules given above for dipole radiation. Other quadripole transitions are found in the L series corresponding to $L-N_{IV}$, $L-O_I$, ... An interesting case is given by Magnusson[1] who finds evidence for N lines in the region tantulum to tellurium corresponding to the transitions $N_{IV}-N_{VI}$ and $N_V-(N_{VI}, N_{VII})$. For these the quantum number n remains unaltered. All these lines are of low intensity compared with the normal emission lines.

Still further sets of emission lines which do not follow from the energy-level diagram are those which are often known as spark lines on account of their analogy to the spark lines in the optical region. In general these lines are in the neighbourhood of strong diagram lines and can be considered as satellites of such lines. Various attempts have been made to explain their origin. A typical argument may be given for the β^{III} and β^{IV} lines in the aluminium and magnesium K spectra. The interpretation of these lines according to Siegbahn and Karlsson is as follows. The first process is the emission of a K electron and an L electron. We can consider this as taking place in two steps. After the emission of the K electron the screening action of the K level is reduced by unity and the action of the nuclear charge on the L electrons is correspondingly increased. The removal of an L electron from the

atom of atomic number Z which has already lost a K electron will therefore require the performance of work approximately the same as that required to excite the L level of an element where nuclear change is $Z+1$. Denoting the double process as $(\overline{KL})$ we have therefore $(\overline{KL})$

$$(\overline{KL}) = K_z + L_{z+1}.$$

If the recombination process is $(\overline{LM})$ we have similarly

$$(\overline{LM}) = L_z + M_{z+1}.$$

For the $K\beta_{\mathrm{I}}$ line we have

$$(\nu/R)_{K\beta_{\mathrm{I}}} = K_z - M_z.$$

For $K\beta_{\mathrm{III}}$

$$\begin{aligned}(\nu/R)_{K\beta_{\mathrm{III}}} &= (\overline{KL}) - (\overline{LM}) \\ &= K_z + L_{z+1} - L_z - M_{z+1}.\end{aligned}$$

Hence for the difference in frequency between the β^{III} and the β_1 lines,

$$\begin{aligned}\Delta\left(\frac{\nu}{R}\right) &= (L-M)_{z+1} - (L-M)_z \\ &\simeq L_{z+1} - L_z.\end{aligned}$$

The frequency-difference calculated in this way is found to agree numerically with the observed difference. Many other lines have been explained on a similar hypothesis, but much work remains to be done before all the observed weak lines and satellites can be considered as satisfactorily accounted for.

The fine structure of absorption edges. Detailed investigations of absorption edges have shown that these are, in general, not simple but possess a fine structure on the short-wave side. This structure is very complex and its interpretation is still a very open question. It seems probable that the fine structure, like many of the satellite lines, arises from double ionization. In the case of zinc, for example, Kievit and Lindsay, who have made very complete experimental studies of the structure of absorption edges, attribute the main K edge to the transition $K - N_{\mathrm{II}}$, N_{II} being the first unoccupied level permitted by the selection rules. The secondary maxima found in the fine structure are considered to be due to double ionization and the first two (eleven in all have been observed) are interpreted as corresponding to $(K-\infty)+(N_{\mathrm{I}}-N_{\mathrm{II}})$ and $(K-\infty)+(N_{\mathrm{I}}-\infty)$. The energies corresponding to these ionizations can be calculated from the data available in the X-ray and optical spectrum measurements and are found to be in good agreement with the observed frequencies.

The intensities of the emission lines. Many measurements have been made of the intensities of the various X-ray emission lines and many attempts have been made to explain the observed intensities and intensity ratios in terms of modern theory. The subject is much too complicated for a full account to be given here, and all that can be said is that although very considerable success has been achieved in the interpretation of the observations much work remains to be done before the question of the relative intensity of the various lines in the spectrum can be considered as placed on a satisfactory basis. A very simple case may be quoted. The $K\alpha_1$, $K\alpha_2$ doublet arising from the transitions $K - L_{\mathrm{III}}$ and $K - L_{\mathrm{II}}$ is such that the intensity

ratio α_1/α_2 is very close to 2 throughout the whole range of the elements. Theoretically this ratio should be determined by the ratio of the values $2j+1$ for the levels. Since j has the values $\frac{3}{2}$ and $\frac{1}{2}$ the calculated intensity is $\frac{2.\frac{3}{2}+1}{2.\frac{1}{2}+1}$, or 2 in agreement with experiment. Similar results are found for other doublets.

The effect of chemical combination. At first it was believed that the X-ray spectra of an atom were entirely independent of its state of chemical combination. The development of measurements of high precision has shown that this is not strictly true, and that the state of chemical combination does have an effect on both the absorption and the emission spectra. This was first detected in the absorption

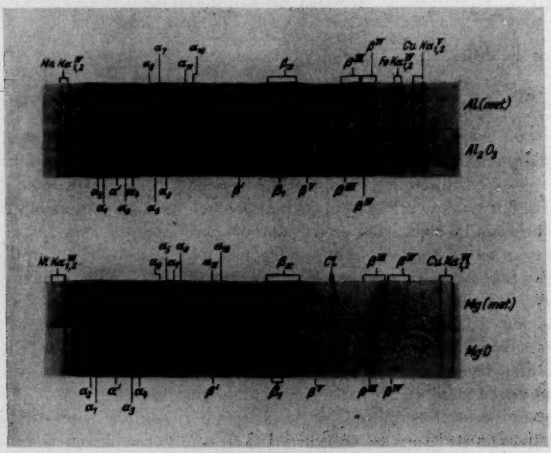


Figure 4. Effect of chemical combination on the the K spectrum lines in the case of Al and Al_2O_3; Mg and MgO.

spectrum by Bergengren and since then many elaborate and beautiful experimental researches have been carried out. The effect on the absorption spectrum is twofold. The actual positions of the absorption edge changes with the state of chemical combination while the fine structure of the edge is also affected. Thus in the case of chlorine there is a gradual shift of the K absorption edge as the valency of the atom changes. The actual wave-lengths of the edge in X. units are 4393·8 for the element, 4382·9 for chlorides (monovalent), 4376·9 for pentavalent compounds such as $NaClO_3$, and 4369·8 for heptavalent chlorates. The secondary fine structure is most complex and covers the widest range of wave-length in the case of the chlorates. Similar results are found for sulphur and phosphorus, and more recent work has shown that the effect is not confined to the light elements but occurs for heavier

atoms such as iron. Corresponding phenomena have been found for the *L* absorption edges.

Later work revealed the fact that not only the absorption spectrum but also the emission lines were affected by the chemical state of the element. Thus Bäcklin found that the $K\alpha$ doublet of sulphur was shifted about 3X. when a sulphate was used as the source of the rays instead of the element. The fine structure and the relative intensities of the satellite lines are also considerably influenced by the state of chemical combination. Thus the $L\alpha$ line for copper or for cuprous oxide appears as a broad diffuse line which is resolved into two sharp components when cupric oxide is used.

In figure 4 is shown a typical example of the effect on the *K* spectrum. The photographs show the *K* spectra of aluminium and magnesium and also of their non-conducting oxides Al_2O_3 and MgO. The $K\beta$ line is of interest. This corresponds to a transition $K-(M_{II}, M_{III})$. The three *M* electrons in aluminium are the valency electrons. In the metallic state the β_x line corresponds to the passage of a conductivity electron to the *K* level. In the metals the β_x line consists of a broad band with a sharp edge on the short-wave side. The intensity increases continuously up to this edge. In the case of the non-conducting oxides a double line (not resolved in the reproduction) β_1 appears displaced towards the long-wave side and characterized by a symmetrical intensity distribution. The breadth of the β_x line for the metals corresponds to a voltage range of 8·4 V. for aluminium and 11·3 for magnesium values which are consistent with those calculated theoretically.

The gap between the X-ray and the optical regions. Until a few years ago a very considerable gap existed between the optical and the X-ray regions. The use of crystals as diffraction gratings for the X rays imposed a very definite upper limit to the wave-lengths which could be measured since it follows from the Bragg equation that these can never exceed twice the largest interplanar distance *d* for the crystal. By the use of crystals with large interplanar spacings such as palmitic acid, it was possible to make measurements up to about 25 A. and Dauvillier actually succeeded in this way in measuring the $K\alpha$ radiation of carbon at 45 A. It was not, however, until new methods of measuring X-ray wave-lengths had been introduced that the gap between the two regions was successfully bridged.

In 1925 A. H. Compton and Doan showed that X-ray spectra could be obtained by the use of ordinary ruled reflection gratings. It will be shown later that in order to obtain results in this way it is necessary to use extremely small glancing angles for the direct X-ray beam since this glancing angle must not exceed the angle of total reflection. For a wave-length of 2 A. total reflection occurs at a glancing angle of 18′ and this increases to 1·5° for a wave-length of 10 A. The necessity of using such small angles introduced many technical difficulties which have gradually been overcome and the more recent measurements obtained in this way show a high standard of accuracy. Such grating methods are particularly suitable for exploring the long wave-length regions and their introduction has led to a great extension of our knowledge of this portion of the X-ray spectrum.

A further advance in the technique of grating measurements has recently been

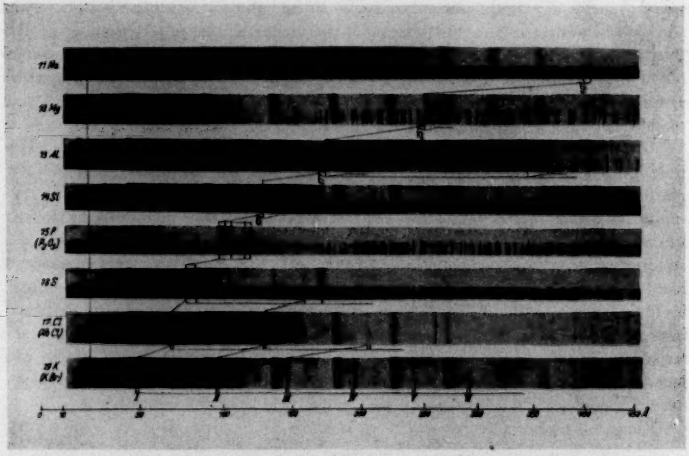


Figure 5. Comparison of the X-ray spectra (L series) of various elements (from Na to K) with the spark spectrum of aluminium.

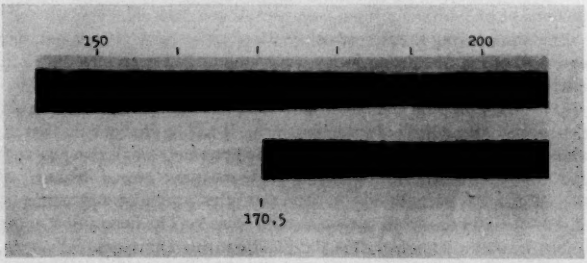


Figure 6. Determination of long-wave absorption edge of aluminium. Upper spectrum: spark spectrum of copper. Lower spectrum: the same spectrum filtered with 0·5-μ. aluminium.

achieved by Siegbahn who has introduced the use of concave instead of plane gratings. In this way the gap between the optical and the X-ray region has been completely filled. In his latest experiments, Siegbahn obtains on the same plate the desired X-ray spectrum and a comparison spark spectrum of aluminium or copper. Accurate measurements of the lines in the spark spectrum which are observed down to about 50 A. are used for the calibration of the X-ray lines. Figure 5 shows a series of photographs taken in this way and illustrating the *L* series from sodium to potassium. Sodium is the element of lowest atomic number for which this series can be obtained, since lighter elements possess no occupied *M* levels. The line spectra on the photographs come from an aluminium spark. The longest X-ray line measured in this region is the *L* line of sodium at 407 A. This line, like many lines of long wave-length, consists of a diffuse band. The strong bands visible on the majority of the plates are successive orders of the $K\alpha$ line of carbon ($\lambda = 43{\cdot}73$ A.).

The concave grating method has also been used for the determination of long wave-length absorption edges. Figure 6 shows a photograph due to Sanner. The upper portion is the spark spectrum obtained from copper electrodes with no filter. For the lower photograph the same source is used but an aluminium filter of thickness $0{\cdot}5\,\mu$. is interposed. There is a sharp cut off of 170·4 A. which corresponds with the L_{III} edge for aluminium. This method offers an accurate alternative method of measuring absorption edges in the extreme long-wave X-ray region.

Absolute measurements of X-ray wave-lengths. Crystal measurements although giving an extremely high accuracy in the determination of wave-lengths relative to the lattice spacing of the crystal depend for their absolute accuracy on the accuracy with which this spacing can be determined. For this reason it is usual to express X-ray wave-lengths measured in this way in terms of X. units where the X. unit, which is approximately 10^{-11} cm., is defined by taking the lattice constant of calcite as 3029·04 X. at 18° C. X.

The use of gratings gives the wave-lengths directly in terms of the metric unit so that by comparison of measurements made in the two ways it should be possible to compare the X. unit with the angstrom unit. Many researches have been undertaken with this end in view. Although the results of such comparisons are perhaps still open to doubt, there seems to be definite evidence that the line-grating method gives slightly higher values for the wave-lengths than those deduced from crystal measurements. Several theories have been put forward to explain the discrepancy. According to one view the results indicate that the generally accepted value of the electronic change *e* is too small by about 0·4 per cent. An entirely different interpretation put forward by Zwicky ascribes the divergence to the essential imperfections which may exist even in the most perfect crystal. Before this important question is settled further research of a high order of accuracy is required.

The continuous spectrum. Among the first studies of X-ray spectra as revealed by crystal diffraction it was found that the complete spectrum comprises discrete lines characteristic of the anticathode material superimposed on a background of continuous radiation. The properties of the continuous spectrum were studied by Duane and Hunt, who showed the existence of a definite short-wave limit at a wave-

length which depends on the exciting voltage and formulated the law that the produce of exciting voltage and limiting wave-length is a constant. Later Hull and Rice established the validity of the law up to 110 kV. and Dessauer and Back showed that it was at least approximately true up to 245 kV., but the difficulties in measuring wave-lengths of 0·05 A. left some slight uncertainty. At lower kilovoltages Wagner showed that the law held down to 7000 V., and Miller, by an absorption method, established its validity to 2500 V. Ledoux-Lebard and Dauvillier showed that, at any particular voltage, the short-wave limit was independent of both the anticathode material and the type of excitation.

The Duane-Hunt law may readily be explained as a result of the quantized nature of the process of producing X rays, the quantum energy of the radiation of shortest wave-length being equal to the energy of the fastest exciting electrons. We have therefore a fundamental relation between the electronic charge (e), the peak voltage on the tube (V), the maximum frequency excited (ν) and Planck's constant (h) in the formula $eV = h\nu$. Observations of the relation between V and ν therefore afford accurate means of determining h. The value of h so determined is

$$h = (6{\cdot}547 \pm 0{\cdot}0009)\ 10^{-27}.$$

An important formula connecting the peak voltage on the X-ray tube and the minimum wave-length (λ_0) emitted is

$$V\lambda_0 = 12{\cdot}330,$$

where V is expressed in kilovolts and λ_0 in angstrom units. The application of this formula affords one of the best methods for determining the peak value of the voltage exciting an X-ray tube.

The distribution of energy in the continuous spectrum was studied by Ulrey and later by Behnken, Dauvillier and Kulenkampff. The energy distribution curve is of the form shown in figure 7, the wave-length of maximum energy (λ_{max}) being about $\frac{3}{2}\lambda_0$. Kulenkampff, who worked at a comparatively low kilovoltage (about 10 kV.), corrected his results for selective action of the diffracting crystal and ionization chamber, and also for absorption in the wall of the X-ray tube and in the anticathode itself, and found that, except near the quantum limit, there is a linear relation between intensity and frequency. Kulenkampff gave an empirical energy distribution formula

$$I_\nu = CZ\,[(\nu_0 - \nu) + bZ],$$

I_ν, ν, Z, C, b — where I_ν is the intensity in the frequency range ν to $(\nu + d\nu)$, Z is the atomic number of the target and C and b are constants.

On account of the difficulties of determining the intensity distribution many workers have studied the properties of the continuous spectrum by determining the curves relating voltage and intensity for any particular wave-length. The determination of such curves (commonly called "isochromats") avoids the uncertainties due to selective reflection by the analysing crystal, absorption, etc. Webster and Hennings showed that the isochromats for a molybdenum anticathode are approximately straight lines except near the critical exciting voltage for the wave-length studied, while Nicolas examined the isochromats of the radiation from a copper

anticathode and showed that their form depends to some extent on the direction of emission.

Some measurements have been made on the total energy emitted by an X-ray tube and its dependence on the anticathode material and the operating conditions. Kaye showed that the total intensity of the radiation from an X-ray tube is approximately proportional to the atomic weight of the anticathode material, a result

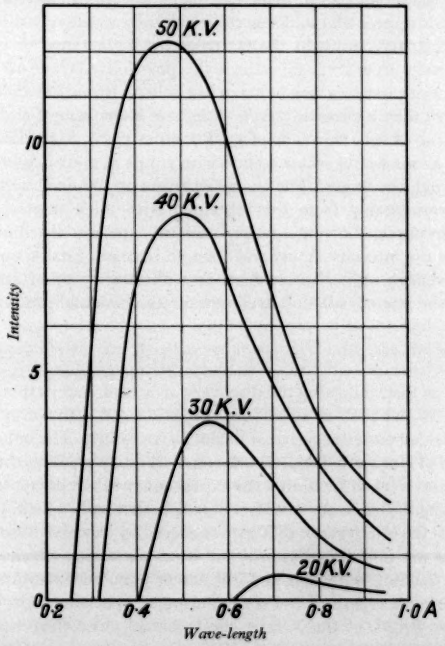


Figure 7. Intensity distribution in the continuous X-ray spectum of tungsten.

subsequently verified by Ulrey and Kulenkampff. Duane and Shimizu studying the radiation from iron, cobalt, nickel and copper showed that the intensity is proportional to atomic number and not strictly to atomic weight, since Co (27) has a higher atomic weight than Ni (28) but gives less X radiation under the same operating conditions.

It has also been established that when full allowance is made for absorption and other disturbing factors, the intensity of the total radiation is proportional to the square of the exciting voltage. This law has been established by Kulenkampff using

an ionization method, Bouwers using a bolometer, Rump by means of an air thermometer and Auren using a thermopile.

The general shape of the energy distribution curve, figure 7, suggests that the emission of the continuous spectrum is governed by some form of probability law, with a limitation on the short-wave side imposed by the quantum process. The long range of wave-lengths may be due to collisions of electrons of a range of velocities with atoms of the target. Even if the impinging stream of electrons is of uniform velocity, it would appear that multiple collisions and intense retarding electric fields within the target may result in the electrons having a range of velocities at the moment of impact. In order to investigate this possibility Duane used as his target a stream of mercury vapour, which being of low density would have a less pronounced disturbing effect than a massive target. Similarly Kulenkampff and Nicholas used targets consisting of very thin foils of gold or aluminium, of thickness about $0{\cdot}6\,\mu$, so that the electrons would either collide with atoms at their maximum velocity or pass right through the target. The energy distribution obtained with targets of this type differs considerably from that obtained with thick targets, the results of Nicholas, for example, showing a nearly uniform intensity distribution from ν_0 to $\nu_0/2$. Similarly the intensity at any wave-length instead of rising continuously and linearly with voltage only rises at first, and afterwards either remains sensibly constant or in the case of radiation in certain directions actually falls with increasing voltage.

§ 7. X-RAY OPTICS

We have seen that, following the discovery of X rays, many attempts were made to detect phenomena such as reflection, refraction and diffraction but that these experiments yielded either negative or inconclusive results. The only exception was the experiment of Haga and Wind, who detected diffraction effects through a narrow slit, but the effects were so small that the experiment was not universally accepted as definitely proving the existence of diffraction. The first truly optical phenomenon to be observed was the interference of X rays scattered by successive atomic planes of a crystal and this was in due course followed by experiments, which established the existence of refraction, reflection and diffraction phenomena similar to those observed in the optical region of the electromagnetic spectrum. On account of the very short wave-length of the X rays, the technical difficulties met with in such experiments are much greater than in the corresponding optical work, but these have been mastered and experiments analogous to many of the classical experiments of physical optics have now been successfully performed for the X-ray region. Of the more outstanding of these we propose to give a very brief account.

The refraction of X rays. The first evidence for the existence of X-ray refraction was obtained in the course of precision wave-length measurements with crystals. It was found that, when the wave-length was determined by measuring the angles through which the X rays were diffracted to give successive orders of the spectrum, the values obtained were not exactly in accord with the Bragg equation,

n, λ, d, θ_n

$$n\lambda = 2d \sin \theta_n .$$

The quantity $\sin\theta_n/n$ was found not to be constant but to diminish as n was increased and only to approach a constant value for high order of n.

If we assume an index of refraction μ equal to $1-\delta$ for the material of the crystal, then a simple calculation shows that, when δ is small,

$$n\lambda = 2d\left[1-\frac{\delta}{\sin^2\theta_n}\right]\sin\theta_n$$
$$= 2d\left[1-\frac{\delta}{\lambda^2}\cdot\frac{4d^2}{n^2}\right]\sin\theta_n.$$

The first of these equations leads to a method of determining δ from the observed values of θ_n. A simple transformation gives

$$\delta = \frac{\sin\theta_1\sin\theta_2(n_2\sin\theta_1 - n_1\sin\theta_2)}{n_2\sin\theta_2 - n_1\sin\theta_1},$$

where θ_2, θ_1 are the observed values of θ for the orders n_1 and n_2. An approximate formula for δ can be calculated from the classical dispersion formula

$$\delta = 1-\mu = \frac{e^2}{2\pi m}\Sigma\frac{N_i}{\nu^2-\nu_i^2},$$

where e and m are the charge and mass of the electron, N_i the number of electrons per cubic centimetre with natural frequency ν_i, and ν the frequency of the incident radiation. For frequencies which are large compared with the natural frequencies, i.e. large compared with the frequencies associated with the various absorption edges of the crystals, this formula becomes

e, m, N_i
ν_i, ν

$$\frac{\delta}{\lambda^2} = \frac{e^2N}{2\pi mc^2},$$

where N is the total number of electrons per cm.^{3} and has the value $\rho Z_M/M.m_H$ in which ρ is the density, M the molecular weight, m_H the mass of the hydrogen atom and Z_M the number of electrons per molecule. From this we obtain for δ/λ^2 the approximate value $1{\cdot}36\times10^{10}\rho$. If these assumptions are justified the value of δ/λ^2 should be approximately constant and should be proportional to the density. As typical of the experimental verification of this formula Larsson found that for a mica crystal and wave-lengths varying from 1·5 to 8 A., δ/λ^2 was very nearly constant and had the mean value $3{\cdot}9\times10^{10}$ (the actual observed values varying between $3{\cdot}6$ and $4{\cdot}1\times10^{10}$), a value which agreed within error limit with that calculated from the density of the mica.

N
ρ, M, m_H
Z_M

Experiments such as these call for a very high order of accuracy and involve the measurements of reflection angles to a fraction of a second of arc. Bergen Davis pointed out that the effect could be considerably amplified if the beam were not symmetrically reflected from a crystal face but if the crystal were so cut that the reflected beam left the crystal at a very small angle. In this way he succeeded in multiplying the effect some thirty times.

Experiments such as those based on the deviations from the Bragg law were capable of a different interpretation. On the basis of the dynamical theory of inter-

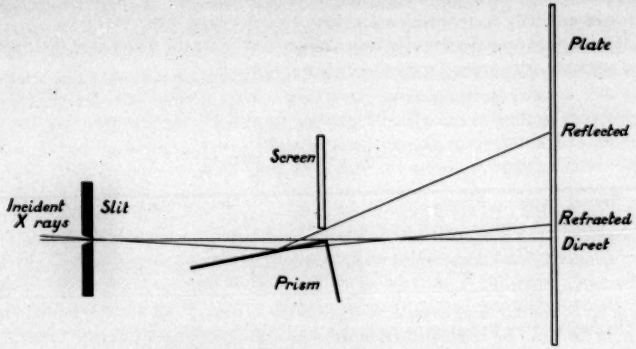


Figure 8. Reflection and refraction of X rays.

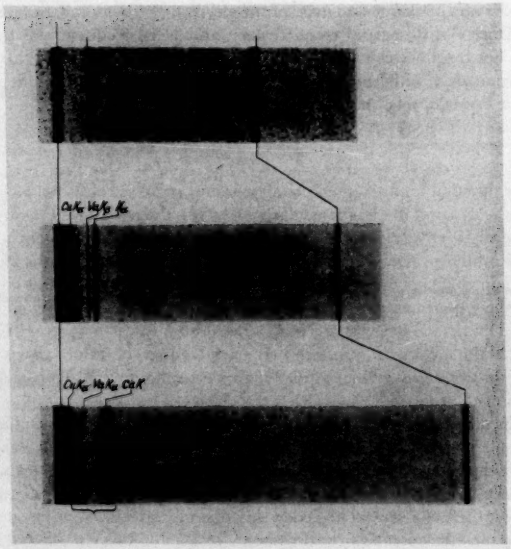


Figure 9. X rays reflected and refracted by a prism. Direct beam on extreme left; reflected beam on extreme right.

ference from a crystal lattice Darwin and Ewald showed that similar deviations should occur and that the formula giving the deviations should be the same as that deduced above on the basis of the refraction of the rays in the crystal. These experiments cannot therefore be considered in themselves as proving the existence of true optical refraction. Many attempts were made to use prisms of amorphous matter such as glass or of polycrystalline materials such as metals, but it was not until 1924 that refraction of X rays was detected and measured in this way. The experiment was first carried out by Siegbahn, Larsson and Waller and consisted in allowing the X-ray beam to fall at almost grazing incidence on the surface of the prism. The direct, reflected and refracted rays were recorded photographically. From these experiments they were able to measure accurately the value of δ for a wide range of materials and wave-lengths. The arrangement is shown diagrammatically in figure 8 and typical photographs are reproduced in figure 9. In these photographs the lines nearest the direct beam on the left are the refracted rays from the quartz prism while the strong lines on the right are produced by regular mirror reflection from the prism surface.

Regular reflection of X rays. It has been shown that refraction experiments led to a value of the refractive index which is less than unity. It follows that total reflection should occur for glancing angles θ where $\cos\theta \geqslant \mu$. Since $\mu = 1-\delta$ where δ is small (of the order 10^{-5} or 10^{-6}) this can be written $\sin\theta \leqslant \sqrt{(2\delta)}$. Owing to the extremely small value of δ it is necessary to use extremely small values of the glancing angle if mirror reflection is to be detected. For glass, for example, $\delta = 1{\cdot}36.\rho.\lambda^2.10^{10}$ so that the glancing angle is given by $\sin\theta = \lambda\sqrt{(2{\cdot}7\rho)} \times 10^5$. Substituting values of λ, we find that even when λ is as great as 10 A., θ has only reached the value $1^\circ\,30'$. It is for this reason that ruled glass line gratings for the measurements of X-ray wave-lengths must be used with very small glancing angles since these have to be kept less than the angle of total reflection. So, too, such gratings are much more suitable for the measurement of long waves than of those in the normal X-ray region since larger glancing angles are permissible with a corresponding gain in accuracy.

Within this glancing angle total reflection occurs and this phenomenon can be used as an alternative method for measuring the value of δ. By observing the maximum glancing angle at which reflection occurs δ can be calculated from the formula $\sin\theta_m = \sqrt{(2\delta)}$. In actual practice the method is only satisfactory when absorption is negligible since only in that case is there a well-defined value of θ_m. As the absorption of the rays in the material of the mirror increases it is found that the reflecting power diminishes rapidly at first and then approximates to a steady value which is about 20 per cent of that found when absorption is negligible.

Dispersion of X rays. Following the development of various methods for determining the refractive index, several researches have been made on the variation of the refractive index with wave-length with especial reference to its behaviour in the neighbourhood of an absorption edge of the refracting material. In the X-ray region a somewhat different law governs the variation with wave-length from that which applies in the optical region. This arises from the fact that for X rays the absorption

of the rays is approximately proportional to the cube of the wave-length. When this is taken into account the theoretical dispersion formula becomes

$$\delta = \frac{e^2}{2\pi m} \sum_i \frac{N_i}{\nu^2} \left(1 + \frac{\nu_i^2}{\nu^2} \log_e \left| \left(1 - \frac{\nu^2}{\nu_i^2}\right) \right| \right),$$

where N_i is the number of electrons per cm.^{3} with characteristic frequency ν_i. If we neglect $(\nu_i/\nu)^2$ for all electrons except the K electrons the formula becomes

$$\frac{\delta}{\lambda^2} = \frac{e^2 N}{2\pi c^2 m} + \frac{e^2 N_K}{2\pi c^2 m} \cdot \frac{\lambda^2}{\lambda_K^2} \cdot \log_e \left| \left(1 - \frac{\lambda_K^2}{\lambda^2}\right) \right|,$$

N_K, λ_K where N is the total number, N_K the number of K electrons per cm.^{3} and λ_K the wave-length corresponding to the absorption edge. This equation is found to agree with observation and to give a reasonably good approximation even in the immediate neighbourhood of an absorption edge. A typical curve showing the anomalous dispersion of calcite is given in figure 10. There is a sudden dip in the value of δ/λ^2

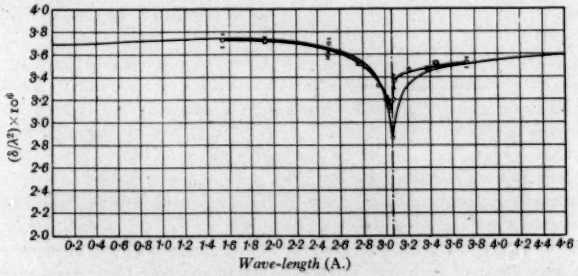


Figure 10. Anomalous dispersion of calcite.

as we approach the wave-length corresponding to the K absorption edge at 3060 X., and a corresponding rise on the long-wave side. The dotted curve gives the variation predicted by the equation given above. Prins has re-calculated the formula taking into account the damping of the electron oscillations and his formula gives a somewhat better agreement with observation.

X-ray diffraction fringes. Diffraction experiments similar to those so well known in the optical field have recently been made in the X-ray region. These include the observation and measurement of the diffraction fringes produced by a Lloyd mirror, the Fresnel double mirror, diffraction round the edge of an opaque obstacle, through a narrow slit, round a fine wire, etc. All of these experiments are of considerable theoretical interest in that they give independent methods of measuring X-ray wave-lengths. On account, however, of the technical difficulties produced by the necessity for extremely fine slits, very small glancing angles and the like, the accuracy of the wave-lengths measured in this way compares unfavourably with that given by crystal and ruled gratings. Apart, however, from their theoretical interest the experiments illustrate so well the refinements possible with modern X-ray technique that a brief account of them must be included in any survey such as this.

In the Lloyd mirror experiment, X rays from a very narrow slit (about $0{\cdot}1\,\mu$) are allowed to fall at an extremely small glancing angle on a mirror, and the diffraction fringes resulting from the interaction of the direct rays from the slit and the rays reflected from the mirror are recorded on a photographic plate set perpendicular to the mirror at a considerable distance from it. If c is the perpendicular distance of the slit from the plane containing the mirror and a the distance of the slit from the plate then successive fringes are separated by a distance given by $a\lambda/2c$. The wave-lengths determined in this way were found to be subject to errors of from 2 to 3 per cent. In figure 11 are reproduced photographs showing typical fringes obtained in this way. These photographs are due to Kellström, who has made a series of very beauti-

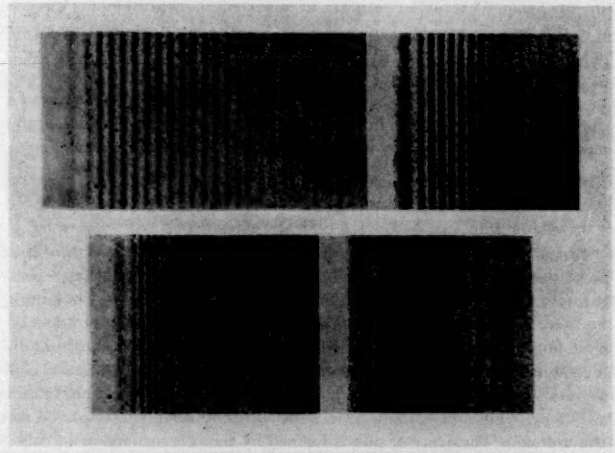


Fig. 11. X-ray interference fringes produced by Lloyd's mirror method.

ful experiments on X-ray fringes. Kellström also used this method to investigate the question of change of phase of the X rays on reflection. For this purpose he obtained on the same plate the fringes due to two wave-lengths of 8·33 and 13·3 A. and examined the coincidences between the positions of the maxima and minima in the two patterns. He concluded that the results were consistent with a phase change during reflection. For the small glancing angles used the phase change amounts to 180°.

Similar fringe patterns were obtained by Kellström by the use of two mirrors inclined at a small angle to one another. In this case the fringes result from the interaction of the beams reflected by the two mirrors. The experiment is exactly analogous to the Fresnel double mirror experiment of ordinary optics, and the results obtained were in accordance with those predicted by the usual theoretical

treatment of the problem. The same worker has also made elaborate studies of diffraction through a narrow slit, past the edge of a half plane and round a thin wire. In the case of the narrow slit he found that the intensity distribution in the fringe system was consistent with the usual formula

$$J = s^2 \frac{\sin^2 \pi xs/a\lambda}{(\pi xs/a\lambda)^2},$$

J, x, s, a where J is the intensity at a distance x from the centre of the pattern, s the breadth of the slit and a the distance separating the slit and the recording photographic plate. Maxima were found to occur at values of x given by $ak_n\lambda/s$ and minima for values $an\lambda/s$, where πk_n ($n = 1, 2, 3, \ldots$) are the roots of the equation $u = \tan u$. From these experiments a somewhat higher accuracy in wave-length measurements was possible so that in this way the relative values of wave-lengths could be determined with an accuracy of 0·5 to 1 per cent. A similar accuracy was obtained from the fringe pattern given by reflection from a small mirror.

The fringes obtained by the bending of X rays past a thin wire or round a half plane have been shown to agree in position and in intensity distribution with those produced by the Fresnel theory and calculated with the help of the Fresnel integrals.

§ 8. APPLICATIONS OF X RAYS

X-ray analysis. As has already been remarked Laue's discovery of the diffraction of X rays by crystals depends on the fact that the atoms in a crystal are repeated at regular intervals, that X rays also are characterized by a repeat or wave-length, and that the two repeats are of the same order of size. Laue's discovery fitted at once into crystallographic theory; it verified the conception of a crystal built by the regular repetition of a unit of pattern consisting of a few atoms. The crystal could be divided into equal parallelopipeds, produced by three intersecting families of parallel planes. One unit of pattern, always similarly arranged, is placed in each cell—the unit cell. The array of points formed by the cell corners, or by any other representative point in each cell, is the space lattice. The lattice points constitute sheets of parallel equidistant planes corresponding to possible crystal faces; each set of planes is defined by the spacing and by the atomic arrangement between any consecutive pair.

W. L. Bragg in 1913 simplified Laue's treatment: he showed that an X-ray beam would be reflected by a set of planes when the glancing angle θ satisfied the relation $2d \sin \theta = n\lambda$, where d is the spacing, λ the wave-length, and n an integer defining the order of the spectrum.

Three experimental dispositions were rapidly developed. In the first, the Laue method, the crystal is stationary and the X-ray beam heterogeneous; certain combinations of plane and wave-length will then satisfy the reflection condition. In the second, adopted by W. H. and W. L. Bragg, the beam is monochromatic and the crystal tilted until a specified plane is satisfactorily inclined. The third utilizes the effect which would arise if the crystal were tilted in every conceivable direction: a

reflected ray would describe, in the course of time, a cone coaxial with the incident beam. The effect is secured, in actual practice, when the single crystal is replaced by a powder of randomly oriented crystallites. The simultaneous reflections from the grains will produce a system of reflection cones, one from each set of planes; they then form a line spectrum on a suitably placed photographic film. This method was used first by Debye and Scherrer in 1916.

The utilization of the spectra involves a study of the six features mentioned in the following paragraphs.

Position of spectral lines. The positions of the lines give the glancing angles, fix the interplanar spacings, and thus indicate the dimensions of the unit cell. The number of molecules in the cell then follows directly from a knowledge of its mass, estimated from the volume and specific gravity, and of the mass of one constituent molecule. Niggli made a further step in 1919 based on the fact that, in general, every order of the spectrum does not appear; restrictions are imposed by the nature of the space group, the combination of symmetry elements associated with the unit cell. This work, developed also by Wykoff and by Astbury and Yardley, led to methods of deducing in a systematic way the space group to which a crystal belongs.

Intensity of spectral lines. The relative intensities of the lines depend on the distribution of atoms between the reflecting planes. The intensities which would result from particular distributions can be computed and compared with the intensities observed. A guide to probable arrangements is given by the atomic and unit cell dimensions, by the number and symmetry of the atoms in the cell, and by chemical and physical properties of the material. The arrangement which satisfies the intensity criteria indicates the atomic structure of the crystal. This application has been developed in particular by W. L. Bragg and his school. The first crystals to be analysed were rock salt and sylvine (W. L. Bragg, 1913). Innumerable structures of increasing complexity have been solved since; a useful summary of the results is published periodically in the *Zeitschrift für Krystallographie* (Strukturberichte).

Structure analysis rendered possible two important applications; to the physics of the crystal lattice, developed first by Born, and to the establishment of a crystal chemistry. Born calculated the energy of a simple lattice; he explained, quantitatively, certain elastic and optical properties, heats of formation and solution, and laid down methods for the study of a system of atoms held in equilibrium by electrostatic forces. This work has been extended by Grimm, Frenkel, Fajans, Lennard Jones and others. The chemistry of the crystal has been developed largely by the work of Goldschmidt and his colleagues. They studied in detail compounds of the form AX, AX_2, AX_3 and A_2X_3 and deduced general laws of structure; crystals were classified in terms of the chemical linkages, and the parts played by atomic size and polarization were discovered.

More technical applications have also been developed. For example, a method of chemical analysis: a monochromatic beam falling on a mixture will produce a composite spectrum in which the spectra of the constituents are superimposed. The

separate spectra, when identified by comparison with known spectra, indicate the components. The method requires only a small amount of material; it gives directly, unlike optical spectroscopy, the state of chemical combination of each constituent; also it permits analysis to be made *in situ* without introducing complications due to chemical manipulation. A further outstanding application is to the structure of the phases of alloy systems. This work, pioneered by Westgren, has resulted in a more precise knowledge of the metallic state; a particular aspect is the discovery of allotropes among which the case of iron forms a classical example. Other applications include the study of chemical reactions, as for example in the nitration of cellulose, and the investigation of the complexes formed by inorganic compounds, as in the case of lithopones.

Breadth of lines. The breadth of the spectral lines, for given experimental conditions, depends on two factors: grain size and lattice distortion. When the grains are less than about 10^{-4} cm., the lines begin to broaden. The effect arises because the resolution of a line is governed by the number of planes co-operating in the reflection. This number, in turn, is fixed by the grain size. Scherrer, Seljakov and Laue have formulated expressions which, when the line breadth is known, permit the estimation of the average grain size. Of these, Laue's formula is the most reliable and general in application. It has been used by Brill in the study of colloidal nickel, and by Wood in the examination of electro-deposited chromium coatings.

Lattice distortion broadens the lines by causing the spacings to deviate irregularly from the normal values. The effect appears particularly in metals and alloys which have undergone deformation by cold working. It has been studied especially in its relation to strain hardening of metals, magnetic coercive force and hysteresis. A special case of distortion is one which produces an abnormal weakening of high order reflections. The phenomenon appears in the plastic deformation of certain metals, in aluminium, for instance, where the line broadening due to distortion is very small.

Continuity of the lines. Each spot making up a line is the reflection from one grain. If the number of grains in an aggregate, exposed to the X-ray beam, is sufficiently small, then the spots do not coalesce. The lines become discontinuous to a degree depending on the number, and therefore on the size, of the grains in the volume illuminated. In ordinary experimental practice, the effect becomes marked as the grains become greater than about 10^{-2} cm. This method of indicating large grain size has been applied particularly to the study of the annealing and recrystallization of metals.

Intensity-distribution along the lines. The intensity along the length of a diffraction line or ring is uniform only if the grains in the aggregate are oriented at random. An uneven intensity distribution will arise as soon as the grains take up a preferred crystallographic orientation. From the position of the maxima and minima of intensity can be calculated the axis of preferential orientation relative to a reference axis in the specimen. Cases of practical importance appear during the fabrication of metals, and have been studied by Polanyi, Schmidt, Wasserman and others. It is found that the operations of wire-drawing, rolling and electro-deposition produce

typical orientation of a nature depending on the metal and the working conditions. Recent work on the changes in the X-ray spectra of plastically deformed metals has led to interesting results. Thus Wood has shown that, beginning with a normal annealed material, the changes occur in a definite sequence; (*a*) the breakdown of the grains to a fixed lower limit of size, usually about 10^{-3} to 10^{-4} cm., (*b*) the production of lattice distortion to an upper limit characteristic of the metal, and (*c*) the occurrence of the orientation effect; and that the physical changes accompanying cold-work take place mainly in the distortion stage.

Preferential orientation in natural substances. A particular aspect of the orientation effect is found in natural fibres. Any process of growth might be expected to result in preferential orientation of the crystals or pseudo-crystals which it produces. Such orientation actually occurs in the case of organic structures such as hair, wool or cellulose, and in the case of inorganic structures such as pearls and the enamel of teeth.

Substances such as hair, wool, cellulose are in a state which is neither crystalline nor amorphous. They are highly polymerized compounds having very complicated molecules composed of a number of similar units (consisting of groups of atoms) which are linked together by chemical bonds to form long molecular fibres. In such substances the fibres are so arranged as to be parallel to each other, i.e. preferential orientation occurs, and there is a repetition of pattern like that in a series of unit cells in a crystal. These pseudo-crystals, besides having a definite repeat of pattern along the direction of the fibres, also possess a regularity of arrangement at right angles to them, owing to the lateral forces between neighbouring fibres. Hence the diffraction effects obtained with X rays resemble those due to real crystals. One of the substances, on which a great deal of work has been done, is cellulose, which consists of polymerized chains of anhydro-β-glucose. A full account of the work is given in a recent book by Meyer and Mark on the structure of highly polymerized substances. Astbury has investigated the proteins and has shown that they can exist in two forms: the contracted form as in unstretched hair, and the expanded form as in silk and stretched hair. Stretched rubber is another pseudo-crystal, consisting of polymerized chains of isoprene and its structure has been investigated by Katz, Hengstenberg and others. From a study of the polyoxy methylenes, Kohlschutter and Sprenger and Sauter have shown precisely how polymerization takes place in this class of substance.

In the case of the inorganic structures, real crystals occur. Thus Shaxby and others have shown that pearls are composed almost entirely of aragonite, crystals being so arranged that their *c*-axes are normal to the surface of the pearl, whilst Thewlis has found that the enamel of teeth is composed largely of apatite, the crystallites of which are so arranged that their hexagonal axes make definite angles with the enamel surface.

Radiography. Radiography has been used in medical diagnosis for a long time. It is now being increasingly applied in industry, to the examination of welds, metal castings and the like for the presence of cracks, flaws, inclusions, etc. Other applications include the examination of coal, rubber tyres, golf balls, insulators, wood and

suspected contraband. In addition, owing to the fact that the old paint pigments consisted of inorganic substances which absorb X rays more heavily than the modern organic dyes, it is possible to examine old paintings and art objects for evidence of retouching, etc.

REFERENCES

Section 1

(1) Taylor Jones. *Induction Coil Theory and Applications* (Pitman, 1932).
(2) Dushman. *Gen. elect. Rev.* **18**, 156 (1915).
(3) Hull. *Gen. elect. Rev.* **19**, 173 (1916).
(4) Bell. *Brit. J. Radiol.* **2**, 156 (1929).
(5) Cockcroft and Walton. *Proc. roy. Soc.* **136**, 619 (1932).
(6) Allibone and Bancroft. *Brit. J. Radiol.* **7**, 65 (1934)
(7) Marx. *Elektrotech. Z.* **45**, 652 (1924).
(8) Allibone, Edwards and McKenzie. *Nature*, Lond., **131**, 129 (1933).
(9) Sarsfield. *J. Instn elect. Engrs*, **67**, 437 (1929).
(10) Coolidge. *Phys. Rev.* **2**, 409 (1913).
(11) Bouwers. *Physica*, Eindhoven, **4**, 174 (1924).
(12) Bouwers. *Strahlentherapie*, **38**, 157 (1930).
(13) Grossmann. *Brit. J. Radiol.* **5**, 252 (1932).
(14) Allibone, Beetlestone and Innes. *Brit. J. Radiol.* **7**, 83 (1934).
(15) Lauritsen and Cassen. *Phys. Rev.* **36**, 988 (1930).

Section 2

(1) *Conduction of Electricity through Gases* by Sir J. J. Thomson and G. P. Thomson, vol. **1**, 1928 and vol. **2**, 1933 (Cambridge University Press) contains a very large number of references relating to ionization in gases. For earlier work, see also J. S. Townsend's *Electricity in Gases* (Oxford, Clarendon Press, 1915).
(2) Anderson, Millikan, Neddermeyer and Pickering. *Phys. Rev.* **45**, 352 (1934).
Blackett and Occhialini. *Proc. roy. Soc.* A, **139**, 699 (1933).
Anderson. *Phys. Rev.* **41**, 409 (1932).

Section 3

(1) Auger. *Ann. Phys.*, Paris, **6**, 183 (1926).
(2) Siegbahn. *Phys. Z.* **15**, 753 (1914).
(3) Richtmyer. *Phys. Rev.* **18**, 13 (1921).
Richtmyer and Warburton. *Phys. Rev.* **22**, 539 (1923).
(4) Allen. *Phys. Rev.* **27**, 266 (1926) and **28**, 907 (1926).
(5) Jönsson. Diss. Upsala (1928).
(6) Backhurst. *Phil. Mag.* **16**, 310 (1933).
(7) Kronig. *Z. Phys.* **70**, 317 (1931), **75**, 191 (1932) and **75**, 468 (1932).
(8) De Broglie. *J. Phys. Radium*, **3**, 33 (1922).
(9) Sir J. J. Thomson and G. P. Thomson. *Conduction of Electricity through Gases* (Cambridge University Press, 1933).
(10) Compton. *Bull. nat. Res. Coun.*, Wash., **4**, No. 20 (1922).
(11) Kramers. *Phil. Mag.* **46**, 836 (1923).
(12) Wentzel. *Z. Phys.* **40**, 574 (1926).

SECTION 4

(1) COMPTON. *Phys. Rev.* **21**, 483 (1923).
(2) SADLER and MESHAM. *Phil. Mag.* **24**, 138 (1912).
(3) MITCHELL. *Phys. Rev.* **33**, 871 (1929).
(4) BEARDEN. *Phys. Rev.* **36**, 791 (1930).
GINGRICH. *Phys. Rev.* **36**, 1050 (1930).
HULUBEI. *Ann. Phys.*, Paris, **1**, 5 (1934).
(5) BACKHURST. *Phil. Mag.* **17**, 321 (1934).
(6) COMPTON. *Phys. Rev.* **35**, 925 (1930).
(7) Sir J. J. THOMSON and G. P. THOMSON. *Conduction of Electricity through Gases*, 2nd ed. p. 325.
(8) DIRAC. *Proc. roy. Soc.* A, **111**, 405 (1926).
(9) DEBYE, BEWILOGUA and EHRHARDT. *Phys. Z.* **30**, 84 (1929).

SECTION 5

(1) NEEFF and REISNER. *Strahlentherapie*, **32**, 181 (1929).
BÖDECKER. *Strahlentherapie*, **47**, 100 (1933).
SCHUBERT. *Strahlentherapie*, **47**, 359 (1933).
(2) BOUWERS. *Z. Phys.* **14**, 374 (1923).
(3) KÜSTNER. *Strahlentherapie*, **24**, 501 (1927).
BEHNKEN. *Strahlentherapie*, **26**, 79 (1927).
FAILLA. *Amer. J. Roentgenol.* **21**, 47 (1929).
KAYE and BINKS. *Brit. J. Radiol.* **2**, 553 (1929), **6**, 530 (1933).
TAYLOR. *Bur. Stand. J. Res.* **2**, 771 (1929).
THORAEUS. *Acta Radiol.*, Stockh., Supplement 15 (1932).
(4) FRICKE and GLASSER. *Z. Phys.* **29**, 374 (1924); *Amer. J. Roentgenol.* **13**, 453 (1925).
JAEGER. *Strahlentherapie*, **30**, 567 (1928); *Phys. Z.* **35**, 184 (1934).
BRAUN and KÜSTNER. *Strahlentherapie*, **32**, 550 and 739 (1929); **33**, 273 and 551 (1929).
MEYER. *Strahlentherapie*, **40**, 576 (1931); **41**, 309 (1931).
KÜSTNER. *Strahlentherapie*, **44**, 759 (1932).
(5) TAYLOR. *Bur. Stand. J. Res.* **8**, 9 and 325 (1932).
KAYE and BINKS. *Brit. J. Radiol.* **6**, 195 (1933).
(6) KAYE and BELL. *J. Instn elect. Engrs*, **73**, 457 (1933).
(7) KAYE. "Protection and working conditions in X-ray departments," *Brit. J. Radiol.* **1**, 295 (1928).
(8) *Protective measures against dangers resulting from the use of Radium, Roentgen and Ultra-violet Rays*, League of Nations Publications, III; *Health*, III. **9** (1931).

SECTION 6

(1) MAGNUSSON. *Z. Phys.* **79**, 161 (1932).

SECTION 7

W. H. BRAGG and W. L. BRAGG. *X-rays and Crystal Structure* (G. Bell and Sons, London, 1924).
SOMMERFELD. *Atombau und Spektrallinien* (F. Vieweg and Sons, Braunschweig, 1922).
SOMMERFELD. *Wellenmechanischer Ergänzungband* (F. Vieweg and Sons, Braunschweig, 1929).
W. L. BRAGG. *The Crystal State* (G. Bell and Sons, London, 1934).
SIEGBAHN. *Spektroskopie der Röntgenstrahlen* (Julius Springer, Berlin, 1931).
COMPTON. *X-rays and Electrons* (Macmillan and Co., London, 1927).
KELLSTRÖM. "Experimentelle Untersuchungen über Interferenz- und Beugungserscheinungen bei langwelligen Röntgenstrahlen," *Nova Acta Soc. Sci. upsal.* ser. IV, **8**, No. 5 (1932).
SIEGBAHN. "Studies in the extreme ultra-violet and the very soft X-ray region" (Guthrie Lecture), *Proc. phys. Soc. Lond.* **45**, 689 (1933).

SECTION 8

W. L. BRAGG. *The Crystal State* (G. Bell and Sons, London, 1934).
NIGGLI. *Geometrische Kristallographie des Diskontinuums* (Borntraeger, Leipzig, 1919).
ASTBURY and YARDLEY. "Tabulated data for the examination of 230 space groups by homogeneous X rays," *Philos. Trans.* A, **224**, 221 (1924).
EWALD and HERMANN. *Strukturberichte*, 1913–28 (Akad. Verlagsgesellschaft, Leipzig, 1931).
BORN. *Atomtheorie des Festen Zustandes* (Teubner, Leipzig, 1923).
JOFFE. *The Physics of Crystals* (McGraw-Hill, New York, 1928).
"Crystal structure and chemical constitution," *Proc. Faraday Soc.* (1929).
Geochemische Verteilungsgesetze der Elemente (Norske Videnskaps.-Akad. Oslo, 1924 *et seq.*).
HUME-ROTHERY. *The Metallic State* (Clarendon Press, Oxford, 1931).
MARK. *Die Verwendung der Röntgenstrahlen in Chemie und Technik* (Barth, Leipzig, 1926).
CLARK. *Applications of X-rays* (McGraw-Hill, New York, 1932).
MEYER and MARK. *Der Aufbau der hochpolymeren organische Naturstoffe* (Leipzig, 1930).
ASTBURY. *Fundamentals of Fibre Structure* (Oxford, 1933).
ST JOHN and ISENBURGER. *Industrial Radiology* (Wiley and Sons, 1934).

THE PHYSICS OF THE ATOM

By Prof. E. N. da C. ANDRADE, D.Sc., Ph.D., F.Inst.P.

MS. received September 15, 1934

PRELIMINARY NOTE

It is obvious, in view of the appalling number of papers that have appeared on atomic structure during the past few years, that in a brief report of this nature only a small fraction can receive reference, and that, a few papers of indisputable importance excepted, the question as to which are to be chosen is a matter for judgment which may well vary widely among individuals. Many elegant pieces of work have been deliberately omitted in this report on account of lack of space. The writer can only offer his apologies to any who may feel that their contributions have not received due prominence, and plead that he has honestly tried to maintain a just balance between the different schools.

Two books of great importance have appeared within the last few years, and on matters that are adequately dealt with in them the writer has felt justified in saying comparatively little. They are *Radiations from Radioactive Substances*, by Rutherford, Chadwick and Ellis (Cambridge University Press, 1930) and *Mass Spectra and Isotopes*, by F. W. Aston (Arnold, 1933). Reference may also be made to the excellent summaries of particular fields of atomic research by K. K. Darrow, issued from time to time by the Bell Telephone Laboratories, in which American work receives an adequate appreciation; the little monographs *Exposés de Physique Atomique Expérimentale*, published under the direction of M. Maurice de Broglie in the series *Actualités Scientifiques et Industrielles* (Paris: Hermann et Cie); and Volume II of the *Comptes Rendus du Congrès International d'Electricité* (Paris: Gauthier-Villars, 1933).

§ 1. INTRODUCTION. GENERAL PRINCIPLES

During the past few years very great advances have been made in the physics of the nucleus. When, at the beginning of 1927, the writer's book *The Structure of the Atom* was published, much had already been done to order the energy levels of the extra-nuclear electrons of the atom, and, although the methods of wave mechanics have recently furnished a satisfactory theoretical explanation of what were then empirical rules, no fundamentally new results have been forthcoming. Since then advances of importance have been made with astonishing rapidity in the physics of the atomic nucleus, and, without disparagement of the other work, it may fairly be said that the main interest of the physicist has shifted to this field.

The discovery of new particles, in particular of the neutron and of the positron, or positive electron, have effected fundamental changes in our outlook. For reasons that will be outlined later we no longer say that the nucleus is built up of protons and electrons, but of protons and neutrons, which is clearly an alternative possibility, since, from the point of view of net charge and (to a first approximation) mass, a proton and an electron in the nucleus are equivalent to a neutron. The artificial disintegrations effected by particles (protons and diplons) accelerated in a discharge tube have greatly extended the earlier results obtained by Rutherford with α-particles and led to the discovery of whole groups of new transformations. The disintegrations produced by neutrons come under the same head. The detailed study of the products of disintegration, especially fast protons, has thrown much light on the question of the mechanics of disintegration. A fresh discovery, of very recent date, is that of artificial radioactivity; it has, namely, been found that the introduction of a new particle within the nucleus, which has been effected by bombardment with neutrons, protons, or α-particles, produces a nucleus which is unstable, and ejects a particle (in nearly all cases an electron, either positive or negative) within a time, of which the average value is, in general, some minutes. In brief, nuclei, of both stable and unstable character, can be artificially manufactured at will from other nuclei.

Of the utmost importance in this connexion have been the studies of Aston, and the one or two other workers who have successfully followed him, on the accurate measurement of nuclear mass. To a first approximation the masses of the individual isotopes, hydrogen excepted, are integral multiples of a number which is 1/16 of the mass of the main oxygen isotope. More refined measurement has shown that there is a departure from this so-called whole number rule, which may be expressed as a "packing-fraction" of the isotope in question. The fundamental importance of the mass of the isotope lies in the interrelation between mass and energy first clearly enunciated by Einstein.

By the generalized theory of relativity mass and energy are mutually equivalent, as are heat and mechanical work by the first law of thermodynamics. In both cases special conditions are necessary to determine the extent to which the interchange can take place, and in which direction it takes place, but there is a constant relationship which governs the conversion, if it takes place. In the case of mass and energy the relationship is given by the equation

$$E = mc^2 \qquad \text{.....(1)},$$

where E is energy, which may appear as a photon or as a particle radiation, m is the mass and c the velocity of light. Thus the apparent annihilation of 1 mg. corresponds to the liberation of 9×10^{17} ergs, or 25000 kilowatt hours.

The consequence is that, if we know accurately the mass of the impinging particle and of the nucleus which forms the starting-point of the reaction, and in addition that of the nucleus formed, we can, from the small difference which usually exists between the two sides of the equation, find the energy with which the particle, or photon, always discharged in these reactions, is emitted. In representing

nuclear reactions it is usual to write the mass number as an upper affix to the chemical symbol of the nucleus in question, and the nuclear charge, or atomic number, as a lower affix; thus $_2^4$He represents an α-particle, $_3^7$Li the heavier of the two lithium isotopes*. The neutron is usually written $_0^1 n$. A typical equation is, therefore,

$$_k^p A + {}_l^q B = {}_m^r C + {}_n^s D \qquad \text{.....(2),}$$

where $$p+q=r+s,$$

and $$k+l=m+n.$$

If the actual masses of the particles in grams† are $p+\delta p$, $q+\delta q$, $r+\delta r$, $s+\delta s$‡, and E_A, E_D are the kinetic energies in ergs, of the impinging particle and of the emitted particle, A and D, respectively, then

$$p+\delta p+q+\delta q+\frac{1}{c^2}E_A = r+\delta r+s+\delta s+\frac{1}{c^2}E_D,$$

$$\delta p+\delta q=\delta r+\delta s+\frac{1}{c^2}(E_D-E_A) \qquad \text{.....(3).}$$

If, therefore, the masses of the various interacting particles, and the velocity, or energy, of the impinging particle, are accurately known, the energy of the emitted particle can be determined, or, *vice versa*, if all the masses but one, and the energy of both the particles, are known, the fourth mass can be determined accurately. This method has been used in arriving at the mass of the neutron. The modification necessary if a photon takes part in the reaction is obvious.

Equation (2) is strongly reminiscent of that for a bimolecular chemical reaction. There seems no reason why an equation for a larger number of participants than two should not be written down, but it is extremely improbable that three nuclei should simultaneously be circumstanced so as to be able to take part in a reaction. This is but one aspect of a general rule, that whereas such an equation as (2) represents a possible reaction, the further question as to the probability of it taking place needs the application of fresh principles for an answer. In the ordinary chemical case we know that the discussion as to how likely a reaction is to take place is given by the equilibrium or reaction constant, which in its turn can be calculated from the heat of reaction ("Wärmetönung") by the help of Nernst's theorem. In the nuclear case the probability of the reaction can be calculated, from energy considerations, by the methods of wave mechanics. In both cases the calculation is possible in principle, but has only been carried out in practice in the simplest cases.

It is not necessary, as it is sometimes suggested, that the mass of the newly formed nucleus shall be less than that of the original nucleus. While this is so in the case of an impinging particle of small energy, if the energy of the particle be

* Common practice is to write the affixes on the right-hand side of the symbol, but I prefer the left, as it leaves place for the ordinary chemical indication of the number of atoms in a molecule—thus $_1^2H_2$ would be a molecule of diplogen (heavy molecular hydrogen).

† It is common practice to express the masses of the atoms in terms of 1/16 of the mass of oxygen. In this case the units of energy must be correspondingly modified.

‡ In this notation the packing-fraction of, e.g., B is $\delta q/q$.

large enough the new mass may clearly be greater than the old. Throughout, the word "particle" is used in the widest sense. It may be a wave-packet or photon.

In working out the results of particle collisions involving the nucleus the conservation of energy and the conservation of momentum are assumed to hold, and this assumption generally proves satisfactory. In certain cases, however, the validity of the conservation of energy has been questioned.

An application of the general principles of wave mechanics to the nucleus has been made by Gamow, with results that are very important for the discussion of many nuclear problems with which we shall be dealing. The essence of the method of wave mechanics is that it allows the probability of various occurrences to be calculated, and it often proves that there is a small but finite probability of an event happening which, on the older views, would be impossible. In much the same way, in classical physics, while geometrical optics says that there will be no intensity inside the geometrical shadow, wave optics enables us to find, in general, a small but finite intensity at any point in the shadow. The probability of the presence of a particle is measured by $|\psi|^2$ or, as it is often written, $\psi\psi^*$, where ψ^* is the imaginary conjugate to $\psi = a + ib$.

The equation, due to Schrödinger, which governs ψ is

$$\nabla^2\psi + \frac{8\pi^2 m}{h^2}(E - V)\psi = 0,$$

where E is the total energy and V the potential energy, so that $E - V$ represents the kinetic energy T of the particle. The solution has an exponential term, and is approximately

$$\psi = \frac{C}{\sqrt[4]{(E-V)}} e^{\pm \frac{2\pi i}{h}\int p\,dx},$$

where $p = \sqrt{\{2m(E-V)\}}$ for a plane wave. The character of the solution clearly depends upon whether $E - V$ is positive or negative; if positive, the exponential has an imaginary exponent, and the solution is a periodic one, in sine and cosine; if negative, the exponent is real, and the solution is exponential. If the kinetic energy, $E - V$, come out to be negative in a region, this would mean in classical mechanics that the particle would not penetrate into that region, but the wave-mechanical solution means that the particle can penetrate, but that the probability of it being found there diminishes rapidly as we proceed into the region, the rate of diminution depending upon the magnitude of $E - V$. If $E - V$ has a large negative value the probability diminishes exceedingly rapidly, and there is practically no chance of penetration, but if the kinetic energy has only a small negative value the penetration has to be taken into account. A somewhat similar consideration arises, in classical wave theory, when total internal reflection is worked out. Besides the reflected wave, there is a wave which is propagated in the less dense medium along the surface of separation, but it is very heavily damped and becomes negligible in a few wave-lengths†.

† E.g. for a wave incident from glass on air at an angle of 60° with the normal the amplitude falls to 1/180 in a single wave-length.

Let us now consider a region bounded by what is called a potential barrier, i.e. a space in which the potential is so high that a particle of given kinetic energy T cannot, on classical reasoning, escape from it. In the upper part of figure 1 the curves $ABCD$, $A'B'C'D'$, symmetrical about the vertical axis, give the potential plotted against the distance r from O. At some distance from O we suppose the field to follow the inverse square law so that the curve approximates closely to the rectangular hyperbola ECD, $E'C'D'$ for moderate values of r: at small values of r the potential drops very rapidly until it has a negative value, approximately constant for still smaller values of r. It seems clear that this represents to a first approximation the field of the nucleus. For simplicity we consider a two-dimensional case, or suppose the potential to have the same value everywhere on a line normal to the plane of the figure.

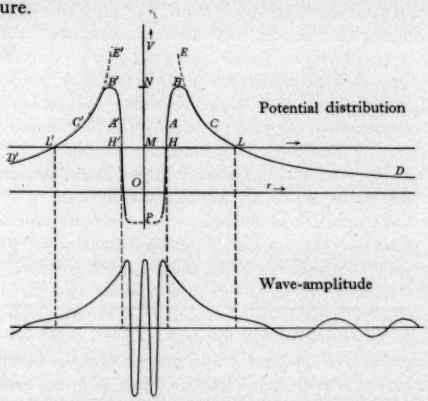


Figure 1.

If PM represents the kinetic energy of a particle on the same scale as the potential energy of the same particle in the nuclear field, then clearly such a particle cannot escape from the space HH': its kinetic energy would be negative in the space HL, $H'L'$. According to wave-mechanical principles, however, the wave, of length $h/\sqrt{\{2m(E-V)\}}$, which represents the particle of kinetic energy T, is mostly reflected to and fro inside the barrier, and takes the form of an approximately stationary wave, but there will be, in addition, a small part inside the region HL, $H'L'$ which decays exponentially as r increases, the decay being extremely rapid if MN is large*. This represents the so-called "leak." A disturbance of small amplitude will reach L, where it enters once more a region where the kinetic energy is positive, so that the exponent in the expression for ψ is imaginary, and we have once more a periodic wave. In the lower part of figure 1 the curve shows the

* No quantitative significance is to be attached to the wave-length shown inside the barrier in figure 1.

amplitude of the wave system; the displacement, of course, fluctuates periodically with time. The wave-length decreases as r increases, owing to the fall in the value of V, and consequent increase in $T=E-V$, until at some distance it becomes $h/\sqrt{(2mE)}$. This decrease in wave-length corresponds, on the particle view, to the increase in kinetic energy as the particle gathers speed in the repulsive field. In the three-dimensional case, which represents more closely the nuclear conditions, the potential barrier taking the form of a symmetrical mountain with a crater in the centre, the amplitude of the wave decreases as r increases, corresponding to the lessened probability, as we proceed outwards, of finding a particle in a given volume.

This is a rough picture of the conditions that govern an α-particle within the nucleus, represented by the potential barrier. The problem was first worked out by Gamow[1,2] and, independently, by Gurney and Condom[3]. The conclusion is that an α-particle, shut in by a region of potential greater than the kinetic energy of the particle, will nevertheless have a finite probability of escaping. A small probability corresponds to a small decay constant. The energy with which the particle emerges corresponds, of course, to the level inside the barrier from which it escapes, and the higher this level the greater the probability of escape, and the greater the kinetic energy which it acquires in proceeding outwards after escape, on account of the large value of V. An element with a small decay constant λ will therefore give an α-particle of smaller energy than one of large decay constant, as is experimentally observed. Gamow has, on general lines, been able to arrive at a quantitative connexion between λ and the energy of the α-particle which agrees as a first approximation with the Geiger-Nuttall relation, derived empirically many years ago. Where the theoretical formula differs from the empirical form the experimental values lean rather towards the former.

Attention should also be directed to a consequence of the fact that the periodic wave within the barrier is reflected to and fro. Just as in the case of a wave on a string fixed at both ends, there is a series of normal modes of vibration, with corresponding definite frequencies. To each frequency ν corresponds an energy $h\nu$, and there will consequently be a series of discrete energy levels for a particle in the nucleus. This conception plays a large part in the discussion of the radiations from the nucleus, both spontaneous and excited.

Another conception, introduced within the last few years, is that of nuclear spin. It is well known that, for the full description of optical spectra, the three original quantum numbers—n, the total quantum number; k, the azimuthal quantum number, or its wave-mechanical equivalent $l=k-1$; j, the inner quantum number—do not suffice. In addition, the electron must have a quantum number s, which may have the value $\frac{1}{2}$ or $-\frac{1}{2}$, and pertains to a vector. This quantum number s is usually explained by supposing that it is due to a spin of the electron about an axis through its centre, $sh/2\pi$ being the angular momentum. This conception of particle spin, associated with a fixed quantity of angular momentum and of energy of rotation, has been applied to the nucleus, which is assumed to possess both angular momentum and magnetic moment. This is further discussed in § 9.

§ 2. NEW EXPERIMENTAL METHODS

The striking experimental advances of the last few years are largely due to new experimental methods which have been evolved. The early information on the nuclear field and on transmutation was mainly obtained by using, as a means of detecting single particles of high energy, the scintillations produced on a zinc sulphide screen and using the α-rays (generally from radium C) to produce disintegrations. The Wilson cloud chamber was also of notable use. Scintillations are now little used as a method of detecting single particles, having been replaced by a counter, used in conjunction with an amplifier based on the properties of thermionic valves. The counter mainly used is a modification of the original Rutherford-Geiger device. It is due to Geiger and Müller[4, 5, 6], and consists of a fine wire arranged axially in a cylindrical chamber. The walls of the chamber may be either thick or thin, according to the radiation to be counted. In the original counter the wire was covered with a thin coat of oxide, which was believed to be essential for the effective working of the detector, but it has been found that practically any wire will serve. In fact, Werner Schulze[7, 8] has pointed out that almost any discharge tube, filled with a gas at a pressure of a few millimetres of mercury and provided with suitable electrodes, can be made to act as a counter, even, for example, a tube with parallel plate electrodes. Nevertheless, there seems little doubt that the cylinder with axial wire is that best adapted for general counting.

By the use of a particularly delicate electrometer, designed by Hoffman[9, 10], it is possible to obtain a direct record of the ionization current produced by a single proton or α-particle, and Pose* has made extensive use of this method. Most workers, however, e.g. Chadwick and his collaborators*, Geiger and his collaborators and Leprince-Ringuet, amplify the current by valve methods, as was first done by Greinacher[11]. The amplified current can be passed through an Einthoven string galvanometer or oscillograph, and a photographic record obtained, in which the deflections are, with suitably designed circuits, proportional to the ionization. Wynn-Williams[12, 13], by the use of the three-electrode valves, with mercury-vapour as the residual gas, known as thyratrons, has devised automatic counters which register the passage of particles on an ordinary mechanical meter, and by ingenious circuits has been able to raise the rate of counting to 1000 a second. In this method, of course, there is nothing to indicate the relative magnitude of the ionization current produced by different particles.

It may be mentioned that Geiger-Müller counters have been extensively used in so-called double and triple coincidence methods. In a double coincidence method two counters are so connected to valve circuits that only if ionization takes place simultaneously in both is a current transmitted[14]. Such an arrangement has been used by Blackett and Occhialini in photographing the cloud tracks produced by cosmic rays†. Triple coincidence, in which ionization must take place simultaneously in three counters, has also been used in cosmic ray research[15], and

* See § 8. † See § 4.

recently a stack of 72 counters has been constructed to register the direction of cosmic rays[16].

The point counter of Geiger consists in principle of a fine rod, usually terminating in a small sphere rather than an actual point, with a plane conductor arranged near and normal to it, which, in the case of particles of small penetration, contains a small window to admit the radiation. Whereas the Geiger-Müller counter responds to rays which transverse it anywhere, the point counter has the particular property, which is often useful, of recording only particles which pass comparatively near the point. It can be arranged, according to voltage, to respond by a trigger action, so that all degrees of ionization within very wide limits produce approximately the same current impulse, or so that the ionization by collision is proportional to the primary ionization[17]. If the potential is adjusted for the latter behaviour, α-particles or protons can be counted in the presence of β-particles, since the latter, on account of the much feebler ionization, are not detected. In both types of counter the sensitiveness is markedly reduced when they are used quantitatively.

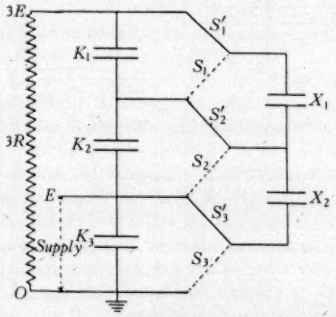
The technique of the Wilson chamber has been so perfected that, in laboratories where swift particles are investigated, it is a standard instrument, just like an electrometer. In fact the electrometer makes rather a good comparison, for just as in the early days of electronic research the setting up and taming of the quadrant electrometer constituted the main part of the work, while to-day electrometers are set up in a morning, so now the Wilson chamber, instead of being a main anxiety, stands more or less ready to hand. Many modifications have, of course, been introduced, to work with gas both at low pressure and at high pressure. C. T. R. Wilson has also designed a new form of chamber, which is very easily constructed. The floor is of wire gauze, through which the expansion takes place. It appears to have certain advantages[18], due to the fact that in it the gas experiences a definite pressure change, instead of a definite volume change, as in the older apparatus.

Another way of counting particles is to use the method of valve amplification in conjunction with a small condenser. The charge produced by the ionization is thrown, by a small potential, on to one of the plates, which is connected to the well-insulated grid of a valve of small inter-electrode capacity, and the consequent small anode current is amplified in the usual way. This method was first used by Greinacher, and has been widely developed. A differential application, due to Rutherford, Ward and Wynn-Williams[19], allows us to measure the range of α-particles with great accuracy, and has been used by them for the long range particles. In this the chamber consists of a thin plane electrode between two other parallel electrodes, one of which is charged to a positive potential and the other to a negative potential. When the ionization is equal in both sides of the chamber, as it is when a swift particle passes through, no charge passes to the central electrode, as the positive and negative ions cancel out. When, however, the particle is near the end of its path, and so losing energy rapidly, the peak ionization will lie in one side or the other of the chamber, with consequent positive or negative charge on the central electrode, for a very small displacement. The method is clearly very sensitive.

Many improvements have been produced in the way of large and constant

magnetic fields, and their wider employment. Mention is made elsewhere of the great magnet at Bellevue, near Paris; of Rutherford's ring magnet; and of the use of the Wilson chamber in a constant magnetic field by Anderson, Blackett and others. For constancy C. D. Ellis[20] has used a permanent magnet of cobalt steel, wound with coils by means of which the field can be altered. When the magnet is in use for an experiment the current is cut off, and the field remains steady for long periods without watching, in contrast to the ordinary electromagnet.

Finally, something may be said as to the generation of high potentials which have played so large a part in artificial disintegration. Many methods have been used, but in most of them a comparatively low voltage is generated in the first place, which is then multiplied by some special device. In Cockcroft and Walton's[21] apparatus the potential generated by a transformer is rectified and multiplied several times by an arrangement of valves and condensers. The general principle is



[*Reproduced by courtesy of the Royal Society.*]

Figure 2.

illustrated in figure 2. K_1, K_2, K_3, and X_1, X_2 are condensers which can be connected alternately either by the full lines or the dotted lines. The primary voltage is applied to K_3, as shown, the switches being in the dotted positions. The condensers X_2 and K_3 are both charged to the full voltage E. The switches are then put over to the full-line position, when K_2 shares its charge with X_2, and so acquires a potential $\frac{1}{2}E$, if the capacities are equal. When the switches pass over to the dotted position again, the condenser X_1 acquires a potential $\frac{1}{4}E$. It is clear that after a large number of alternations the condenser K_1 will be charged to the full potential E, and that the total potential applied to the end of the high resistance, which represents a discharge tube, will be $3E$. In practice the switching takes place by means of hot-filament valves of special construction. For details the reader must be referred to the original paper.

E. O. Lawrence and M. S. Livingstone[22] have perfected a very ingenious method in which the multiplication is much higher, so that the primary voltage can be quite low. The fundamental fact on which it is based is that the time taken by an

ion to describe a circular path in a uniform magnetic field is independent of the speed, the radius of the wider circle described being proportional to the velocity. The apparatus consists of a flat circular metal box, cut into two by a diametral gap, as shown in figure 3 (*a*). By means of a water-cooled power tube oscillator an alternating potential difference of some 4000 volts peak value can be applied to the two halves of the box, which is placed between the plane pole pieces of a very large magnet, producing a uniform field of some 14,000 gauss. A source of ions is arranged near the centre of the box: it is formed by a stream of electrons from a hot filament which passes normally to the flat faces of the box, through a small hole in the side, and ionizes the residual gases. The ions are accelerated by the difference of potential between the two halves of the box, and describe a small semicircular path. The frequency of the alternating field is so arranged that as the ions issue

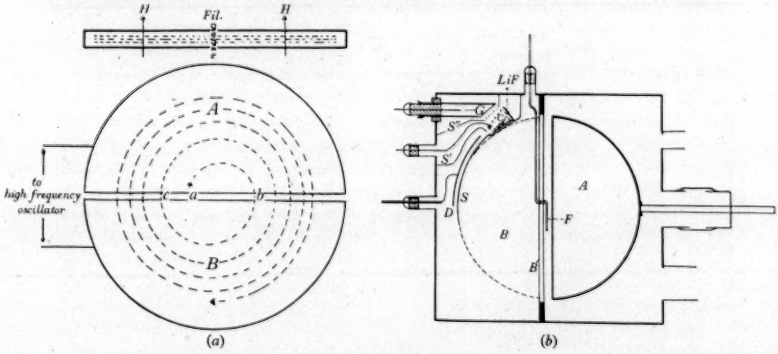


Figure 3.

from the one half box they are once more accelerated into the other half, the potential having just changed sign. Owing to the extra velocity they describe a somewhat larger semicircle, and on issuing are again accelerated into the other half box, in consequence of the isochronic property mentioned above. According to Henderson, who has used the apparatus in Lawrence's laboratory to investigate the disintegration of lithium under proton bombardment, as many as 300 impulses may be given to the ions before they pass out of the accelerating chamber. With the field and potential above specified, and a frequency corresponding to a wave-length of 14 m. he obtained a final radius of 11·5 cm. and an energy corresponding to a potential of 1,123,000 volts. Figure 3 (*b*) shows the arrangement used by Henderson: *A* is the accelerating electrode, the other half *B* being earthed. *F* is the filament for producing the ionizing electrons, *D* is a deflecting electrode which draws the protons after their final circuit out into a circular path which ends in the collecting box *C*. This contains the target, while *G* is a Geiger counter, separated from the target by the usual thin foils.

§ 3. THE WORK ON ISOTOPES

The first stage of the work on isotopes may be said to have been completed round about 1924, when Aston had definitely shown that in all cases where the atomic weight differed markedly from a whole number the element consists of isotopes whose masses are very close to whole numbers. The establishment of the "whole number rule" allowed us to assume that all nuclei are built up of the same elementary particles. Since that date a few new isotopes have been discovered, some, notably those of oxygen and hydrogen, of remarkable importance, but the main attention has been devoted to the question of establishing the exact masses of the isotopes, in other words, to determining the small amount by which the masses depart from being whole numbers. This is but one more version of the old physical paradox; first the establishment of constants, and then further research to see how the constants vary.

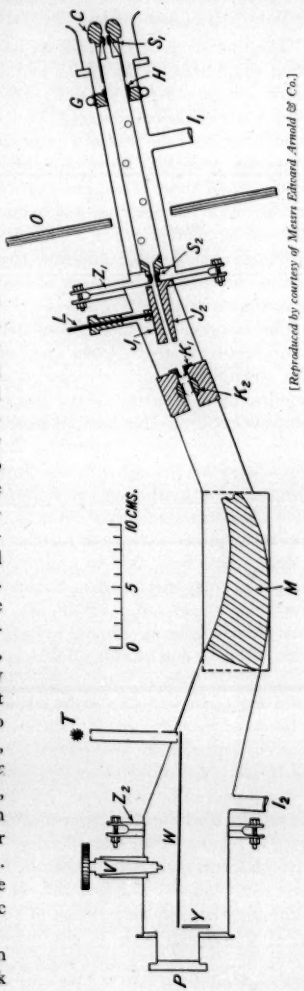
The accurate determination of the mass of isotopes has been effected mainly by two means: firstly by Aston's new mass spectrograph[23], built on the lines of his older instrument, but improved so as to give five times the resolving power; secondly, by the instrument of Bainbridge, originally designed as an improved Dempster apparatus, but later modified so as to depart considerably from Dempster's pattern. Aston's new instrument, illustrated in figure 4, had an improved slit system, in which the two slits, S_1, S_2, 0·02 mm. across, were less than half as wide as the earlier apparatus, and were twice as far apart; an improved electric field, in which the plates were curved, so as to follow the mean deflected ray, which enables them to be placed very close together, so that a comparatively low potential will produce a high field; and an improved magnet, giving a very constant field. The pole pieces were machined to the special shape shown at M. The photographic plate, about 16 cm. long, is at W, the rays being turned through about 10°. As the actual electric fields were found to be below that given by the voltage applied to the plates, owing to polarization effects caused by some kind of adsorbed surface layer, which are not constant, various special methods have been devised by Aston which do not depend for accuracy on the assumption that the applied potential difference is operative between the plates. One generally used by him is carried out as follows. A line corresponding to known mass is brought very near to the line of the unknown mass by using two different applied potentials, whose ratio is approximately that of the two masses being compared. The dispersion curve of the instrument, giving the relation between linear separation and mass, can be found by calibration to a sufficient degree of accuracy over a small range* to enable the unknown mass to be found with high accuracy, supposing that the ratio of the fields effective between the plates were exactly known. This is found not from the measured potential difference, but by choosing lines corresponding to known masses, whose ratio is approximately that of the known mass to the unknown mass in the first experiment, and,

* The dispersion constant can be found to within about 0·3 per cent, so that, if the interval corresponds to a difference of mass of 1 per cent, an error not exceeding 1 in 30,000 may be anticipated in this connexion.

using the same applied potentials, changing the magnetic field so as to bring these two known lines into the same region of the plate. In both cases the two electric fields are applied alternately throughout the experiment, the relative duration of the two being arranged so that the intensities of the two near lines are about equal. Aston's estimate, which seems to be a conservative one, of the accuracy obtained by such a method is 1 in 10,000.

Aston has further developed a photographic technique for determining the relative abundance of the isotopes of a single chemical element. We may recall that although some workers—Dempster, Hindley and G. P. Thomson—had originally found that the relative abundance of e.g. ^{6}Li and ^{7}Li depended upon the conditions of experiment, it has been clearly established by Aston, Morand and Bainbridge that the ratio is a constant one, indicating that in the original formation of the isotopes an equilibrium state was reached. Aston's measurements are carried out by the use of special plates, and a microphotometric routine. Clearly if the masses of the isotopes are accurately known, and their relative abundance known to within a per cent or two, the atomic weight can be very accurately found, supposing that the less abundant isotopes are not too distant in mass or too important. To make this point more precise, Aston introduces the conception *isotopic moment*, which is the sum of the product, relative abundance, by difference between mass of isotope and mean atomic mass, for all the isotopes of the element in question. Clearly, the larger the sum the more the secondary isotopes influence the atomic weight.

The important results reached by Aston in the second development of his work are:



[Reproduced by courtesy of Messrs Edward Arnold & Co.]

Figure 4.

(1) The determination of the exact masses of individual isotopes, which, as has been pointed out, is of the utmost importance in considering the energy relations in nuclear disintegration.

(2) The determination of atomic weights, which can be found very accurately, from the masses and relative abundance of the isotopes. Indeed, in many cases Aston's figures have led to chemical redeterminations, which have always been favourable to the isotopic method, while in some cases Aston's value is accepted as the best available. Obvious advantages of Aston's method are the small quantities required, and the fact that the purity of the specimen is of no importance, atoms of impurity being sorted out in the tube.

(3) The settlement of certain radioactive questions, where Aston's determinations of the isotopes of lead have been of great importance. The lead isotope 207, found by him in a pure specimen of uranium lead, can only have come from the actinium series, since uranium cannot give an odd mass number, while in ordinary lead 208, which is very faint in Aston's uranium lead, is much stronger than 207. This settles the question of the atomic weight of all the numbers of the actinium series, hitherto uncertain: the atomic weight of protactinium, from which lead is formed by six α-changes, is 231. This indicates that the parent of protactinium, atomic weight 235, is an isotope of uranium, called by Rutherford actino-uranium. Taking the ratio of the number of 207 atoms to 206 atoms as 7 to 100, which agrees with Aston's latest determination, Rutherford deduces as the half-life period of actino-uranium $4{\cdot}2 \times 10^8$ years. Supposing that the uranium in the earth was originally formed in the sun, before the earth was separated from it, and has been decaying since, it can then be calculated that the time which has elapsed since that separation is $3{\cdot}4 \times 10^8$ years.

While the old parabola method of J. J. Thomson has been elaborated, with the production of astonishingly fine curves by Eisenhut and Conrad[24] and by Lukanow and Schütze[25], with the idea of studying the chemical reactions in discharge tubes, and while Smythe and Mattauch have devised an apparatus on a new plan which has been used for the oxygen isotopes, to be discussed, the only worker besides Aston who has systematically produced results of a high order of accuracy with several elements is Bainbridge. He took as his starting-point Dempster's method, in which the ions, released practically without energy, are given a known energy in an accelerating field, and analysed in a constant magnetic field. He used a particularly large and steady magnetic field, the pole pieces being some 40 cm. across, which permits of high resolution. He soon abandoned Dempster's method of producing homogeneous rays for a new one, using what he calls a "velocity selector." The deflection, over a given distance, of a beam of ions of given e/m and of velocity v, in an electric field X is proportional to X/v^2, while in a magnetic field H it is proportional to H/v. If the actions of the two fields are opposed, it is clear that for the deflections to cancel, the velocity must be proportional to X/H, and that for any other velocity, X and H being fixed, there will be a finite deflection. If, therefore, a bundle of rays, of all velocities, generated in a discharge tube in the ordinary way, be isolated by two slits S_1 and S_2, and then passed through a space

across which the two fields prevail, only ions whose velocity is in the neighbourhood of a fixed value will pass through the third slit S_3. This is the principle of the velocity selector; the rays leaving it are separated, according to the value of e/m, in the uniform semicircular magnetic field, and detected by a photographic plate. In the earlier method of Bainbridge they were, as in Dempster's original experiment, detected by a Faraday cylinder and electrometer. This latter method is particularly satisfactory for precise measurements of the relative abundance of isotopes, and was used by Bainbridge for the ratio of ^{6}Li to ^{7}Li, as already mentioned. It should be pointed out that the "velocity selector" transmits constant velocity, whereas Dempster's method of acceleration in a known field gives, for a given charge, constant energy. This has certain interesting consequences, discussed in Aston's book. The beam strikes the plate normally, as can be seen from figure 5,

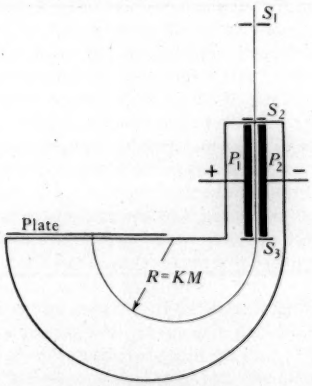


Figure 5.

and in consequence the lines are very sharp and symmetrical. The dispersion is a little greater than that of Aston's instrument. Bainbridge has made several extremely accurate determinations of mass, especially of the hydrogen-helium mass ratio, for both H_1 and H_2. The very close agreement between the results of Bainbridge and Aston is particularly satisfactory in view of the completely different types of apparatus.

In the past few years extensive use has been made of other methods of demonstrating the existence of isotopes, methods which are not adapted to give precise measurements of isotopic mass, but rather to reveal the existence of isotopes present to the extent of a very small fraction only of the main isotope. These methods are all spectroscopic. In the case of line spectra certain features of the hyperfine structure can be explained on the basis of differences of nuclear mass in chemically identical atoms. This is mentioned in the section on the spin of the nucleus, where reference is made to the discovery of the lead isotope ^{204}Pb.

The most striking discoveries in the field of isotopes during the past few years have been the new isotopes of oxygen and hydrogen. Since oxygen as 16 has been adopted as the standard of atomic mass it is very important to know if ordinary chemical oxygen contained any isotope beside the ordinary ^{16}O. Aston devoted careful attention to the point, and concluded from his experiments that, if there were such an isotope, it could not exist in proportions greater than 1 in 500. About the same time W. F. Giauque and H. L. Johnston announced, as a result of a study of the absorption band spectrum of the oxygen of the atmosphere, that isotopes ^{17}O and ^{18}O existed which manifest themselves as the molecules $^{16}O\,^{17}O$ and $^{16}O\,^{18}O$. Actually ^{17}O had already been detected in the Wilson chamber, since Blackett had photographed a nitrogen nucleus which, struck by an α-particle, had captured that particle and emitted a proton. Giauque and Johnston estimated the proportions as 1 in 1250 of ^{18}O and 1 in 10,000 of ^{17}O, but the later work of R. Mecke and W. H. J. Childs has amended these figures to 1 in 630 and 1 in 3150, which are still outside the limit given by Aston, so that they do not in any way invalidate his work.

Now the standard in all the work of the mass spectrograph is the line given by the ^{16}O isotope, while the standard in any chemical work is the three isotopes mixed in the proportions just given. With these values for the abundance, the atomic weight of ordinary oxygen is $16{\cdot}0035 \pm 0{\cdot}0003$ in terms of the atomic weight of ^{16}O as 16. This means a difference of 2·2 parts in 10,000 between atomic weights taken with reference to the mass spectrographic standard and those measured chemically. All Aston's atomic weights need, then, to be reduced by this amount to bring them to the chemical scale, which is just in excess of the limit of experimental error.

For hydrogen the chemical atomic weight is $1{\cdot}00777 \pm 0{\cdot}00002$, while Aston's best determination, with reference to ^{16}O as 16, was 1·00778. When the oxygen isotopes were discovered, this latter value had to be corrected to 1·00756 on the chemical scale, so that instead of the exact agreement previously thought to exist, there was a discrepancy between the mass-spectrographic and the chemical value which was greater than the estimated error of experiment. At the same time, the difference, as Aston[26] has pointed out, is only just on the border of experimental uncertainty, since there are three independent experiments involved, the chemical atomic weight, the abundance of the rare isotopes of oxygen and the isotopic weight, and Aston's estimate of the possible error in the last named was 1·5 parts in 10,000. Nevertheless, the discrepancy was sufficient to encourage Urey in collaboration with Brickwedde and Murphy[27] to look for a heavy isotope of hydrogen, a very small fraction of which would suffice to explain it. From theoretical considerations they came to the conclusion that the molecules $^1_1H_2$, $^1_1H^2_1H$, $^1_1H^3_1H$ should have different vapour pressures in the liquid state, supposing the heavier isotopes to exist. They proceeded to distil liquid hydrogen fractionally under different conditions of pressure, and examined spectroscopically the gas coming over in the last stage. In the case of a heavy isotope the Balmer lines of the atomic spectrum of hydrogen should be shifted slightly towards the short wave-length end of the

spectrum, owing to the effect of the finite mass of the nucleus, of which the lines of ionized helium afford a familiar example. For $^{2}_{1}H$ the shift is about 2·7 parts in 10,000. With samples of hydrogen evaporated at low pressure in the neighbourhood of the triple point the displaced lines could be clearly detected for all four lines H_α, H_β, H_γ and H_δ, the measured displacement agreeing well with the value calculated for $^{2}_{1}H$.

Since then efficient methods of separating $^{2}_{1}H$ from ordinary water have been evolved. Washburn and Urey[28] first discovered that the acid in old electrolytic cells was particularly rich in heavy water, and G. N. Lewis and R. T. Macdonald[29,30] have worked out a method of concentration by electrolysis which enables them to prepare heavy water, of density 1·11, in practically a pure state. It appears that $^{2}_{1}H$ normally exists in hydrogen (and, of course, in hydrogen compounds) to the extent of about 1 part in 6500. There is no space here to refer to the physical and chemical properties of the new isotope, which have been summarized in a variety of publications[31,32,33]. From the point of view of the study of nuclear mechanics the new atom has proved of great importance. Accounts of the effect of bombardment by diplons* on light nuclei will be found elsewhere.

It may be mentioned that the experiments on transmutation have given evidence for certain new isotopes, the existence of which has hitherto not been confirmed in any other way. Prominent among these are the hydrogen isotope, $^{3}_{1}H$, and the helium isotope $^{3}_{2}He$, whose temporary existence, at any rate, has been made probable by the experiments of Oliphant, Harteck and Rutherford, which receive reference elsewhere.

It may be added that these and other nuclei have been demanded by the builders of nuclei on paper. A large number of such schemes have appeared recently. They are difficult to summarize in such a way as to do full justice to the authors, and are so generally accessible that the reader will forgive me if I content myself with this reference.

§ 4. THE POSITIVE ELECTRON

The positive electron was discovered as a result of investigations on cosmic rays by means of the Wilson cloud chamber. The chamber is placed in a strong magnetic field, as was first done, apparently, by Skobelzyn[34], so that the tracks of particles are curved; the magnitude of the curvature depends on the mass and velocity of the particle, supposedly carrying unit charge, and on the strength of the field, while the direction of the curvature in a given field depends on the sign of the charge and the sense of the velocity vector. The mass of the particle producing a cloud track can be deduced by considering that, while it is involved in the curvature, or, better speaking, the $H\rho$, of the track, it does not affect the ionization, which depends upon the charge and velocity only. In a word, two particles with the same

* Rutherford, Discussion on Heavy Hydrogen, *Proc. roy. Soc.* **144**, 1 (1934), has suggested the names "diplogen" for heavy hydrogen, and "diplon" for the nucleus of the new isotope, which have met with wide acceptance. The name "deuterium," proposed by Urey, with "deuton" for the nucleus is also used—see references [31,32,33].

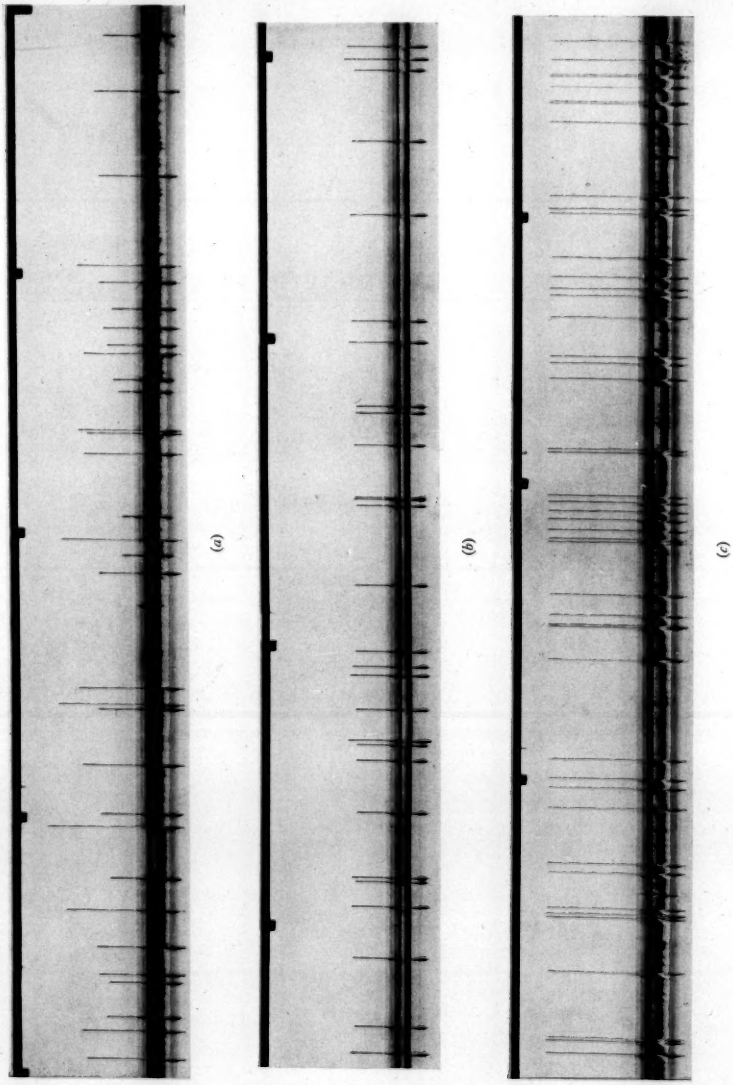

(a)
(b)
(c)

PLATE I

Records of particles made by an oscillograph connected to the output of a valve amplifier used in conjunction with a small condenser (see p. 276). The time signals at the top indicate intervals of 10 seconds. I am indebted to Dr J. Chadwick for these records, made especially for this report.

(*a*) Protons.

(*b*) α-particles, with low sensitivity.

(*c*) α-particles, with higher sensitivity. Every kick goes off the record.

PLATE II

The positive electron (p. 285)

(*d*) and (*e*) Pair of photographs showing about 16 separate tracks produced by a cosmic ray. The divergence is from a point in the copper coils of the magnet, some way above the top of the photograph. The track curved markedly to the right must be due to a positive electron.

(*f*) A particle which has passed through a 4 mm. lead plate. The curvature is much greater below than above the plate. The direction of travel must therefore be downwards, which makes the curvature correspond to a positive charge. (There is a change of direction in passing through the plate which is irrelevant for our purpose.)

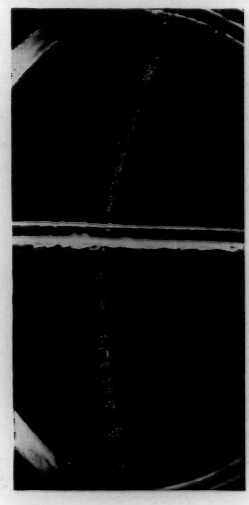
[(*d*), (*e*) and (*f*) are from Blackett and Occhialini, *Proc. roy. Soc.* A, **139**, 699 (1933), by permission of the authors and of the Royal Society.]

Artificial disintegration (p. 295)

(*g*) Two separate disintegrations of lithium by accelerated protons. The tracks c_1, c_2 and d_1, d_2 are due to opposite pairs of α-particles.

[From P. I. Dee and and E. T. S. Walton, *Proc. roy. Soc.* A, **141**, 733 (1933), by permission of the authors and of the Royal Society.]

(*d*) (*e*)



(*f*)

(*g*)

[*Reproduced by the courtesy of the Royal Society*

charge and the same $H\rho$, but different masses, will produce different intensities of ionization, this being, for ordinary values of $H\rho$, more intense for the greater mass, which implies a smaller velocity.

When a cosmic ray passes through or near the chamber a number of tracks are produced, which correspond to a mass of the order of that of the electron, some being curved in one direction and some in the other, which in itself might be due to differences in the sense of the velocity. Anderson[35,36,37,38,39], in his early experiments, obtained a decisive track, which traversed a lead plate some 6 mm. in thickness. The curvature was much greater on the one side of the plate than on the other, thus bearing evidence to a diminution of velocity which must indicate that the electron had already traversed the plate. The sense, so fixed, of the velocity vector showed that the charge of the particle must be positive, and Anderson stated specifically that this, and one or two other tracks which were also indicative of a positive charge, must be due to particles with a mass very much less than that of the proton. Meanwhile Blackett and Occhialini[40] were working on similar lines, but their apparatus had the great advantage that it was automatically released by a unit of cosmic radiation traversing it, so that a photograph was only taken when there was something to photograph. As a result tracks occur on 80 per cent of Blackett and Occhialini's photographs, while Anderson obtained only 2 per cent with tracks suitable for measurement, and Skobelzyn[41] not more than 10 per cent. The expansion was initiated by a device controlled by two Geiger-Müller counters, arranged one above and one below the Wilson chamber, so that any ray which traversed them both also went through the illuminated part of the chamber. By suitable circuits, and the use of a thyratron, it was arranged that only when there is coincidence of ionization impulses in the two counters, which in general means that they have been actuated by the same cosmic particle, does the chamber function. The expansion is completed within 0·01 second of the discharge of the counters, and the illumination takes place within 0·01 second of the expansion. The ions have thus not time to diffuse away sufficiently to make the tracks inconveniently broad.

Blackett and Occhialini, like Anderson, used the difference of curvature on the two sides of a plate traversed by the track as an indication of the sense of the velocity, but they also observed another phenomenon which gave valuable information on this point. In many cases the photographs showed groups of tracks, up to twenty or so, which radiate from a point outside the chamber, presumably an atom which, from geometrical considerations, may form part either of the air or of the various metals in the neighbourhood. In such "showers," as the experimenters call them, the sense of travel is known, since on no reasonable hypothesis can it be supposed that several independent particles move simultaneously in to one point, especially all within a limited angle. These showers generally contain particles curved both to the right and to the left, of approximately equal range, which is much greater than would be possible if the particle were a proton. The density of ionization also indicates a mass very much too small for the proton and comparable to that of the electron. There seems, then, to be no doubt as to the

existence of the positive electron. According to Anderson it has been established by counts of ions that the mass of the positron is equal to that of the electron within 20 per cent, and the charge equal to that of the electron within 10 per cent. Reference to a more accurate measurement by Chadwick and others will be made a little later.

The question arises as to how it is that the positron was not detected earlier. In this connexion calculations of Dirac, cited by Blackett and Occhialini[42], are of particular interest. He has proved that the life of the positron must be very short, depending on the concentration of negative electrons, but of the order of 10^{-10} seconds in water, for instance, and, of course, longer in air. The life predicted by Dirac is long enough to enable the positron to be observed, just after generation, in a cloud chamber, but short enough to explain why it has never been detected by other methods. If a wave radiation of large $h\nu$, reacting on matter, can produce a pair of electrons of opposite sign, then we have for the kinetic energy E,

$$E_- + E_+ = h\nu - c^2 (m_- + m_+) \qquad \ldots\ldots(4),$$

since the energy equivalent of the created mass m is mc^2.

Positrons have now been produced by other agencies than the cosmic radiation, in particular by the high-energy γ-rays from Thorium C″ as found by Anderson[43], the Joliots[44], Meitner and Philipp[45] and Chadwick, Blackett and Occhialini[46]. The latter experimenters, using γ-rays for which the maximum value of $h\nu$ is $2{\cdot}62 \times 10^6$ ϵ volts*, have been able, by measuring the maximum energy of positive and negative electrons, as obtained from the cloud tracks of some hundreds of particles, to compare the masses of the electron and positron. It is clear from equation (4) that the maximum value of the energy of the positron is given by the right-hand side, since E_- may be negligible. The energy equivalent of the mass of the electron is 511,000 ϵ volts, so that if $m_- = m_+$ then the maximum value of

$$E_+ = h\nu - 1{\cdot}02 \times 10^6 = 1{\cdot}60 \times 10^6 \ \epsilon \text{ volts},$$

which agrees closely with the experimental result. Considering also the maximum energy of the electrons, released photoelectrically from the K shell of lead, they are able to compare the masses of electron and positron, and find

$$m_+ = (1{\cdot}02 \pm 0{\cdot}10)\, m_-.$$

The maximum values were used in these calculations because a pair of electron tracks is seldom obtained, no doubt due to the absorption of one or the other in the lead used as the source in the experiments. The hypothesis that a positron-and-electron pair is produced simultaneously is, however, at the basis of the calculations, which are confirmed by experiment. Another conclusion that can be drawn from the energy measurement is that the pair is produced in the extra-nuclear part of the atom rather than in the nucleus. On this view the presence of the nucleus merely plays the part of a catalyser. It must be stated, however, that Anderson considers some of his results to show that the nucleus plays a direct part in the release of positrons by γ-rays.

ϵ volts

* The expression "ϵ volts" denotes electron-volts, i.e. the energy which an electron would acquire moving freely through the potential fall given.

In any case, general energy considerations show that a photon cannot produce a positron and an electron in free space. If $k=\frac{h\nu}{c}$ is the momentum of the photon and p^+, p^- are the momentum of positron and electron respectively it can be proved that $|\vec{k}-\vec{p}_+-\vec{p}_-|$ must be > 0. In free space there is nothing to take up the surplus momentum, but in atomic conversion the momentum can be transferred to the atom or constituent of the atom. The question of the probability of the conversion, with which the theoretical men are now employed, is a difficult problem in wave mechanics.

§ 5. THE NEUTRON

The discovery of the neutron, a particle with the mass of the proton but carrying no charge, was the result of work in Germany, France and England, but the critical experiments found Cambridge ready to recognize their implication, since, many years ago, Rutherford[47] had contemplated the possible existence of a particle possessing mass but no charge. He had even looked for it, but without success. The first step towards the new discovery was furnished by the experiments of Bothe and H. Becker[48,49,50], who in 1930 were working on the effect of bombarding various elements with the α-rays from polonium, which have the advantage of being free from the accompaniment of β- and γ-rays. They were looking for long-range protons on the lines of the experiments of Rutherford, Chadwick, Pose and others. They found that certain light elements, notably lithium, boron and fluorine, gave rays which passed through 2 mm. of brass, while beryllium was particularly productive of such rays. The rays were more penetrating than was to be anticipated of any known corpuscular radiation, and the experimenters assumed without question that they were γ-rays. It may be said at once that the existence of such γ-rays has since been confirmed, but the point is that they are not the only radiation produced under these conditions.

The work was continued by Joliot and Irène Curie-Joliot[51], who used a much stronger source of polonium and showed that the radiation excited in the beryllium nuclei could penetrate several centimetres of lead; since then it has been detected through 30 cm. of this metal. The French workers found that the beryllium radiation could expel protons of high energy from paraffin wax, protons which they detected in a Wilson cloud chamber with a thin window. They also found that other light nuclei could be projected by the radiation. The range of these protons was, however, such as to offer great difficulties on the supposition that the radiation was of a wave nature, communicating momentum to the protons by the Compton effect. It was left for Chadwick[52], who, within a day or two of the French publication, was experimenting on the subject, to point out that all the difficulties could be met if the radiation was of a particle nature.

Both a wave-packet and a particle can transfer energy and momentum to a second particle. In both cases a particle of a certain energy, $h\nu$ for a wave-packet, and approximately $\frac{1}{2}mv^2$ for a particle, strikes the proton, say, considered to be at

rest; the result of the collision is that the proton moves off in a certain direction with a part of the energy, while the impinging particle moves off in another direction with the remainder. In the case of an impinging particle, however, the momentum is obtained from the energy by multiplying by 2 and dividing by the velocity of the particle in question (supposing that relativity considerations can be neglected), while in the case of the wave-packet we divide by c. Since we have both energy and momentum equations, the laws of collision are clearly different in the two cases, one where we deal with particles only, and the other where we deal with wave-packets and a particle. This affects both the maximum energy that can be communicated to an impelled nucleus by a corpuscle (photon or particle) of given energy, and the way in which this maximum range varies with the mass of the impelled nucleus*.

Chadwick[53] argued that the impelling agent could not be a photon, firstly from the number of impelled protons, which can be theoretically deduced for high energy quanta from the Klein-Nishina formula, and secondly from their range. There are far too many protons, and the range would demand a photon of about 55×10^6 ϵ volts energy. There is not enough surplus energy in the formation of a ^{13}C nucleus from a ^{9}B nucleus and a captured α-particle to liberate this. He then compared experimentally the maximum range of protons impelled from paraffin wax by the new radiation with that of other impelled nuclei, and found that the relative values were inconsistent with a photon hypothesis. On the other hand, all the experimental results could be consistently explained by supposing that the radiation produced from the beryllium by α-particle bombardment was a particle radiation. The mass of the particle had to be about that of the proton to satisfy the collision relations. The lack of charge had to be assumed to account for the great penetrating power of the radiation, and was consistent with reasonable transformation equations, such as

$$^{11}_{5}B + ^{4}_{2}He = ^{14}_{7}N + ^{1}_{0}n,$$

which represents the birth of a neutron. At the same time, Feather[54] examined the collision of the neutrons with nitrogen nuclei in the Wilson chamber and obtained direct proof of disintegration. The neutron gives no track, on account of its lack of charge and consequent lack of ionizing power, so that the result of the collision appears as an isolated event in the gas filling the chamber. The lengths and angles of the short tracks of struck nitrogen nuclei which appear on the photographs agree with the neutron hypothesis. The tracks are sometimes single, corresponding to elastic collision; sometimes they consist of two short tracks, making an angle with one another, due to a nucleus formed by capture of the neutron subsequently disintegrating, with liberation of an α-particle. There is also evidence of dis-

* It is common to apply the expression "recoil atom" or "recoil nucleus" to the particles set in motion by impact in nuclear collisions of all types. The term, however, seems to me unfortunate. The word "recoil," in ordinary parlance, suggests the "set-back" of some body which takes place when it discharges a projectile in the opposite direction, and it was in this sense that it was applied to atoms in the early days of radioactivity: when an atom discharged an α-particle the main (transformed) nucleus moved in the opposite direction as a recoil atom. To-day the term is applied to a particle struck by another particle in an elastic collision. It is suggested that "impelled atom" is a more appropriate term, as this may be considered to include both elastic collision, and the case of capture and subsequent discharge.

integration without capture of the neutron. It may be added that the fact that impelled protons never make an angle of more than 90° with the estimated path of the neutron is also evidence for its particle nature.

The penetrating power of the neutron is very great. Reference has just been made to the lack of interaction with matter, consequent upon the lack of charge. Dee has shown experimentally that a neutron does not produce more than 1 pair of ions in a path of 3 m. in air, while a charged particle produces at least 12,000 in this distance. Work carried out by M. de Broglie and Leprince-Ringuet, Thibaud[55, 56, 57] and others in his laboratory has shown that the interposition of 30 cm. of lead leaves 10 per cent of the incident radiation unabsorbed, but that certain light substances have a preferential absorption. For instance, 6 cm. of paraffin reduce the radiation by 50 per cent, a reduction which requires 8 cm. of lead, and aluminium too has a relatively high absorption. Fermi[58] has suggested that this may be due to a kind of resonance effect, the de Broglie waves for the neutron being of the order 10^{-12} cm. in length, which is of the same order as the sphere of action of the forces between neutrons and nuclei. To illustrate his point he cites the Ramsauer effect. It is clear that before the question of transmission through matter can be satisfactorily settled the true absorption of the neutron, such as occurs in Feather's transmutation experiments, must be carefully distinguished from sideways scattering, either elastic or inelastic.

As regards the γ-rays originally found by Bothe and Becker, Meitner and Philipp[59] have not only confirmed their production when beryllium and other light elements are bombarded by α-rays, as Rasetti[60] has also done, but they have shown that, with beryllium, boron and lithium, the ratio of the intensity of γ radiation to neutron radiation is about the same from element to element. According to I. Curie and Joliot, γ-rays are always found as an accompaniment of neutron emission. The energy of the γ-ray from beryllium has been measured in particular, and amounts to about 5×10^6 ϵ volts[60, 61]. It is supposed, taking this element as an example, that in the reaction

$$^{9}_{4}\mathrm{Be} + ^{4}_{2}\mathrm{He} = ^{12}_{6}\mathrm{C} + ^{1}_{0}n,$$

the carbon nucleus is formed in general in an excited state, and gives out the γ-ray on its return to the ground state*. The measurement of the range of neutrons shows that there are at least two groups of different energy: Chadwick[62], for instance, finds, for beryllium, neutrons of velocity $2{\cdot}8 \times 10^9$ and $4{\cdot}7 \times 10^9$ cm.sec., the latter being fewer in number. It is suggested that the carbon nucleus is occasionally formed in the ground state, when a more energetic neutron is emitted: in general it is formed in the excited state, with emission of a less energetic neutron, the balance of energy being emitted as a γ-ray when the nucleus returns to the ground state.

The question of the mass of the neutron has been studied by several workers. Chadwick deduced a value from the reaction

$$^{11}_{5}\mathrm{B} + ^{4}_{2}\mathrm{He} = ^{14}_{7}\mathrm{N} + ^{1}_{0}n,$$

* See § 7.

the energy of the neutron being deduced from the maximum measured range, and the energy of the nuclei from Aston's mass determinations. This gave 1·0067 as the mass of the neutron. Curie and Joliot, however, believe that the boron reaction may be

$$^{10}_{5}\mathrm{B} + ^{4}_{2}\mathrm{He} = ^{13}_{6}\mathrm{C} + ^{1}_{0}n + \dot{\epsilon},$$

the symbol $\dot{\epsilon}$ denoting the positron*, or in other words Chadwick regards the neutron as a proton + electron, and Curie-Joliot the proton as a neutron + positron. They also envisage other possible methods of transformation of the ^{10}B nucleus. Their considerations give a mass 1·012. The mass has been deduced from other reactions: the average value of several collected by R. Ladenburg[63] is 1·0062, omitting an exceptional value obtained from diplogen. The neutron has also been obtained as a product when lithium is bombarded by high speed diplons, and the reaction

$$^{7}_{3}\mathrm{Li} + ^{2}_{1}\mathrm{H} = 2\,^{4}_{2}\mathrm{He} + ^{1}_{0}n$$

should give the mass of the neutron with high accuracy[64]. The value found is 1·0062. The question of the mass is a difficult one, since every energy change involved in the reaction from which it is deduced must be known to a high order of accuracy. It seems that Chadwick's original value cannot be far from correct, and it is unlikely that the mass can lie outside the limits 1·006 to 1·010.

Many other questions, such as the dependence of the range of the neutron on the energy of the releasing α-particle, have been diligently investigated, but their discussion would take us too far afield. It may be taken that the existence of a neutron, with no charge and a mass closely approximating to that of the proton, is well established, and that this particle plays a part in many transmutations.

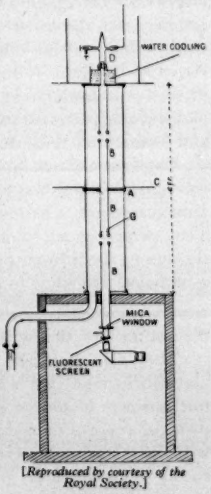
§ 6. ARTIFICIAL TRANSMUTATION

In the radioactive substances we have a spontaneous transmutation of one chemical element into another; the term artificial transmutation is applied to cases where such a transformation is provoked by the action of particles of high energy, including in the term "particles" wave quanta, should transformation by their agency prove to be well established. Artificial transmutation was first described by Rutherford just after the War, when he showed that α-particles could expel protons from the nuclei of nitrogen atoms, a result which he soon extended to other light elements, or, more precisely, to elements of low atomic number, since it is the small nuclear charge which conditions the close approach of the α-particles to the bombarded nuclei. It is essential that the impinging particles should have high energy if they are to be effective in producing transmutation, and in this respect the α-particles are particularly indicated: for the particles of 7 cm. range from radium C, for instance, the energy is about $7{\cdot}7 \times 10^{6}\,\epsilon$ volts. Other advantages possessed by the α-particles are the compactness and easy manipulation of the minute source, and the uniformity of the energy, nearly all the particles having the same range. It is probably correct to say that the most precise quantitative work on nuclear trans-

* See § 4.

formation is still carried out with α-particles, and a glance through this report will show how great a part they played in recent discoveries.

At the same time, they have the drawback that only a comparatively small number is produced by an ordinary source. A milligram of radium, or any α-ray product in equilibrium with this amount of radium, emits some $3{\cdot}7 \times 10^7$ particles a second, which may sound no inconsiderable number until it is remembered that only about 1 particle in 10^6 produces a transmutation under typical conditions, owing to the very small impact parameter which is effective. For comparison with figures to be cited later it may be noted that the α-ray emission from 1 mg. of radium corresponds to a current of about $1{\cdot}2 \times 10^{-11}$ amperes. The difficulty of preparing a pure source and the rapid decay of the best α-ray sources are other familiar difficulties. Again, the direct radiations from the radioactive substances provide us with only one type of projectile of sufficient energy, the helium nucleus. Accordingly, Rutherford soon realized and emphasized the importance of devising a method for the artificial production of a copious stream of particles of high energy. The acceleration of ionized or stripped atoms in a discharge tube is the obvious method of obtaining such particles, but it was met by the equally obvious difficulty that a potential drop of a few million volts appeared, on the prevailing views, to be necessary if disintegration was to be effected. The production of such high voltages is within the range of present possibilities, but there remained the difficulty of building a tube to stand them.



[*Reproduced by courtesy of the Royal Society.*]

Figure 6.

Fortunately, however, it has proved that far lower voltages are sufficient. Transmutation involves, in general, the penetration of an incident particle into the nucleus, and here we are confronted with considerations of the penetration of a potential barrier, similar to those which arise in the discussion of the emission of α-particles from a radioactive nucleus. The impinging particle may emerge from the field of the struck nucleus without loss of energy (elastic collision), as in the experiments of Rutherford and his collaborators which established a departure from Coulomb's law in the case of intimate collisions. On the other hand, there may be provoked energy interchanges with the nucleus, leading to the emission of other particles (inelastic collision). The point which concerns us immediately is that particles with an energy far less than that sufficient to penetrate the potential barrier, on the older views, may, as shown by the principles of wave mechanics, still have a finite chance of penetration, as pointed out in § 1. If, therefore, the particles from our artificial source are very plentiful, we may anticipate transmutations at potentials far below those originally contemplated.

The earliest successful transmutations with a non-radioactive source of particles

were carried out by Cockcroft and Walton[65, 66], with potentials of a few hundred thousand volts. The method used by them to obtain these potentials has already received reference. The ions to be accelerated were generated in a hydrogen discharge tube of ordinary Wien type, which was excited by a small alternator sufficiently well insulated from the floor to stand the 800 kV. contemplated, and driven from the floor by an insulating belt. The acceleration of the ions took place in the tube shown in figure 6, built up of the wide outer glass tubes, separated by a metal plate *C* maintained at half the total potential, to distribute the electric stress. The inner axial tubes *BB* are of steel and are provided with thick rings at the end to hinder discharge. They serve to focus the ions, which are generated in the tube *D*.

If experiments are to be carried out at atmospheric pressure the protons emerge through a thin mica window: otherwise a target may be placed in the brass chamber which forms the end of the vacuum tube, and which is provided with a suitable window or windows. In the earliest experiments the particles thrown off by the target were detected by the scintillation method, but this was soon replaced by a counter, with the usual amplifier and oscillograph. The Wilson chamber was also used, with a very thin window of mica to admit the particle, in order to check ranges. Recently Dee and Walton[67] have used the Wilson chamber to investigate the method of transmutation of lithium and boron, but, needing a more copious stream of disrupting particles, they have had recourse to a tube of the type used by Oliphant and Rutherford, to be described shortly.

The first element to be systematically investigated by Cockcroft and Walton was lithium, which was exposed to a proton stream representing a current of up to 5 microamperes. The result was an emission of α-particles, the range being about 8 cm., which agrees on energetic grounds with the interpretation that the particles are due to the lithium isotope ${}^{7}_{3}\mathrm{Li}$ capturing a proton and breaking up into two α-particles, according to the equation

$$ {}^{7}_{3}\mathrm{Li} + {}^{1}_{1}\mathrm{H} = 2\,{}^{4}_{2}\mathrm{He} + 16{\cdot}8 \times 10^{6}\,\epsilon \text{ volts} \qquad \ldots\ldots(5), $$

the last term on the right representing the kinetic energy of the particles.

They also observed, using as detector an ionization chamber with amplifier and oscillograph, the way in which the number of particles per unit time per microampere of proton stream current varies with the energy of the impinging protons, a type of measurement which has been extensively carried out in connexion with disintegration experiments. It is of great interest in connexion with Gamow's theory of the nucleus, giving as it does the probability W of a particle of given energy entering a nucleus. If Ze is the charge of the particle, of mass m and energy E, while the charge of the struck nucleus is $Z'e$, then

$$ W = e^{-AZZ'e^{2}/\sqrt{E}}, $$

where $A = \frac{1}{h}\,4\pi\sqrt{2m}\,J_k$, J_k being a function which varies slowly with E and Z. Cockcroft and Walton worked up to 500 kV., disintegrations being still observed with energies as low as 70 kV. The results were speedily extended by the Berkeley

school of investigators—by Lawrence, Livingstone and White[68] up to 710 kV., and by M. C. Henderson[69] up to 1125 kV. There is a very rapid increase of number of α's produced with voltage, there being, for instance, about 10 times as many with 1200 kV. as with 400 kV. The general run of the curve agrees with Gamow's theory, but the fact that all the experiments were carried out with thick targets, so that protons of energies much below the maximum were operative as well as those of maximum energy, renders comparison somewhat difficult. Experiments with thin targets are described a little later.

Cockcroft and Walton established the disintegration by proton bombardment of several other light elements, especially boron, fluorine and aluminium, with liberation of α-particles. These are elements which, some years ago, Rutherford and Chadwick transformed by α-particle bombardment, with liberation of protons. They also disintegrated lithium, boron and carbon by bombardment with diplons—ions of heavy hydrogen[70]. This important class of missiles was first used for disintegration in the country of its discovery by Lawrence, Lewis and Livingstone[71]. The American workers, using particles of from 600 to 1330 ϵ kV., found that with a number of elements they obtained protons of the same range, viz. about 18 cm., a result which they explained by supposing that the diplon itself was disrupted into a proton and a neutron. They also announced other proton groups, and, with certain elements, an emission of α-particles. Cockcroft and Walton found with the elements mentioned above that long-range protons (30 cm. with lithium, and smaller ranges with the others) are emitted on diplon bombardment, and from energy considerations they conclude that the bombarded nucleus captures the neutron of the diplon, thus transforming itself into the next heavier isotope, while the remaining proton is emitted. Thus for lithium

$$^{6}_{3}\mathrm{Li} + {}^{2}_{1}\mathrm{H} = {}^{7}_{3}\mathrm{Li} + {}^{1}_{1}\mathrm{H},$$

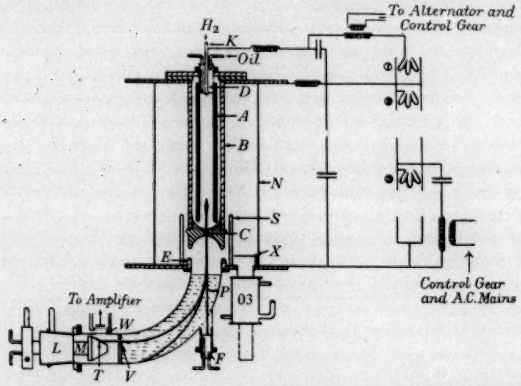
the energy and mass changes giving $5 \times 10^{6}\,\epsilon$ volts as the energy of the emitted proton. With lithium and boron there is also an α-particle emission, of about the same frequency of occurrence; thus

$$^{10}_{5}\mathrm{B} + {}^{2}_{1}\mathrm{H} = 3\,{}^{4}_{2}\mathrm{He}.$$

This transformation has been particularly investigated by Rutherford and his collaborators, in work to be described shortly.

As soon as Cockcroft and Walton's results had clearly shown that low voltages (500 kV. is here called a low voltage, since it is so in comparison with the millions of volts originally contemplated as necessary) were adequate for transmutation, provided that the impinging particles were sufficiently numerous, Oliphant and Rutherford[72] set to work to construct an apparatus to produce a much heavier stream of charged particles at a voltage not to exceed 250 kV. They also arranged for the ions to pass through a uniform magnetic field of up to 18,000 gauss, which serves both to separate out ions of different masses and to ensure that the energy is the same for all the ions, of one kind, used for transmutation. The abundance of bombarding particles enables very thin sheets to be used as targets (stopping power of the order 1 mm. of air), so that a proton never makes more than one collision with

a nucleus. The result is that the energy of collision is the full energy of the bombarding particle, which is uniform and accurately known, so that curves can be obtained which show the true relation between energy of impinging particle, expressed in kilovolts, and number of particles emitted. The apparatus is represented in figure 7. The ions are generated within the space enclosed by the steel cylinder *B* and its end-piece *C*, the latter being pierced by a hole through which the ion stream passes. The enclosure made up of *B* and *C* acts as the cathode, the anode consisting of the steel cylinder *A*, this unusual construction being effective in diminishing the pernicious sputtering and heating effects consequent on the heavy discharge currents involved. The accelerating field is applied between *C* and *E*. Heavy pumps



[*Reproduced by courtesy of the Royal Society.*]

Figure 7.

Figure 7 (*a*).

maintain the pressure in the space traversed by the emergent ions at a much lower pressure than that prevailing in the discharge tube, this maintenance of a difference of pressure on two sides of an orifice having been first introduced by Wien, and made generally familiar by its use in the mass spectrograph. A proton stream carrying a current of the order of a milliampere has been thus produced.

With this apparatus, used by Rutherford in collaboration with Oliphant and Kinsey[73], light elements have been bombarded with copious streams of monoenergetic* protons and diplons, the latter from a specimen of heavy water presented by G. N. Lewis. Lithium exposed to proton bombardment gives α-particles of range 8·4 cm., as found by Cockcroft and Walton, and also of two short ranges, of 1·15 and 0·65 cm. respectively. Diplon bombardment produces protons of even longer range, viz. 13 cm., corresponding to an energy of $11{\cdot}5 \times 10^6\,\epsilon$ volts, and a continuous group containing all ranges from 1 to 7·8 cm. The long-range α-particles

* I have ventured to coin the word "monoenergetic" for a stream of particles having all the same energy, by analogy from monochromatic.

obtained with protons can be attributed to the reaction of equation (5) which checks up well with the masses and energies involved. The short ranges are more difficult to explain. As regards diplon bombardment, Lewis, Livingstone and Lawrence[74] had already obtained long-range α's from lithium, and attributed them to two α's expelled in opposite directions, with, of course, equal energies, according to the equation

$$^{6}_{3}\mathrm{Li} + {}^{2}_{1}\mathrm{H} = 2\,{}^{4}_{2}\mathrm{He}.$$

Lewis, Livingstone and Lawrence find a range of 14·8 cm., while Rutherford and co-workers find 13·0 cm., which corresponds better with the energy relations. The continuous group is attributed by Rutherford and his colleagues to the reaction

$$^{7}_{3}\mathrm{Li} + {}^{2}_{1}\mathrm{H} = 2\,{}^{4}_{2}\mathrm{He} + {}^{1}_{0}n.$$

If the three particles are emitted simultaneously, as seems likely, the simplest assumption is to suppose that the two α's are emitted symmetrically with respect to the neutron, as shown in figure 7*a*. The fraction of the total energy E possessed by the individual particles will be a function of θ. If θ is 0° the energy of each α-particle will clearly be about 1/16th of that of the neutron, the respective values being 0·056 E and 0·889 E, while if θ is 90° the energy of each α is 0·5 E, and that of the neutron is zero. The distribution obtained by taking into account all the other possible angles agrees with the experimental results[75]. A similar consideration of the simultaneous emission of three particles has been applied to the case of the disruption of the boron isotope $^{11}_{5}\mathrm{B}$ by the impact of protons, with the production of three α-particles[76]. On the assumption that the number of particles emitted corresponding to any angle θ falls off exponentially with the difference between this angle and the symmetrical value 120°, very good agreement with the experimental curve of the absorption of the mixed beam of α-particles was obtained.

The products of artificial disruption have also been investigated by the method of the Wilson chamber, first by Kirchner[77] and then by others, in particular by Dee and Walton[78] who have devised special types of windows allowing particles leaving the target in any direction to pass into the surrounding expansion chamber. They have confirmed the splitting of $^{7}\mathrm{Li}$ into two α's by proton bombardment, the range, 8·3 cm., agreeing extremely well with that derived from energy considerations and the latest data for the masses of the nuclei, and have further confirmed the mode of disintegration obtained by the English and American workers for lithium under diplon bombardment, and boron under proton bombardment. The beautiful pictures give particularly striking evidence of the complex transmutations which are being produced.

A remarkable result obtained by Oliphant, Harteck and Rutherford[79] relates to the effect of bombarding diplons with diplons. Whereas the α-radiation from polonium does not produce any effect on diplogen, diplons of 100 ϵ kV. energy provoke a heavy radiation, consisting of a group of protons of range 14·3 cm. and a group of singly charged particles of range 1·6 cm., the numbers of each kind of particle being identical. The experimenters attribute the short range particles to a hitherto unknown isotope of hydrogen, $^{3}_{1}\mathrm{H}$, the reaction being

$$^{2}_{1}\mathrm{H} + {}^{2}_{1}\mathrm{H} = {}^{3}_{1}\mathrm{H} + {}^{1}_{1}\mathrm{H}.$$

It may be mentioned that Latimer and Young[80] believe that they have shown the existence of ${}^{3}_{1}\mathrm{H}$ in heavy water, by a magneto-optic method, but the amounts must be excessively minute, probably not more than 1 to 5×10^{8} in ordinary water. In further experiments Rutherford[81] and his collaborators have also detected a neutron radiation, which they attribute to a second possible mode of disintegration

$$ {}^{2}_{1}\mathrm{H} + {}^{2}_{1}\mathrm{H} = {}^{3}_{2}\mathrm{He} + {}^{1}_{0}n, $$

thus involving a new isotope of helium, ${}^{3}\mathrm{He}$. Experiments with a Wilson chamber carried out by Dee[82] serve to confirm this.

Besides the proton, diplon and helium nucleus, the neutron has been used as an agent to provoke transmutations. The transformations have to be observed by the Wilson chamber, since, beside the range, the direction of the track or tracks must be known if the nature of the collision is to be determined. Feather[83] bombarded nitrogen gas with neutrons produced from beryllium irradiated by α-particles. Some 1500 pairs of photographs were taken, and a number of tracks starting in the gas were recorded, some single and some paired, i.e. consisting of two tracks starting from the same point. Analysis showed that the first were to be attributed to the liberation of a proton by the neutron, which was not itself captured, while the second cause is due to the capture of the neutron, and consequent projection of the modified nucleus, while at the same time an α-particle is liberated. Similar photographs have been secured by Harkins and his colleagues at Chicago[84].

It is pretty certain, too, that cosmic rays can effect atomic transmutations, as the bursts of particles observed by Blackett and Occhialini and by Anderson can hardly be produced without nuclear disintegration.

Only a few of the most important transformations have received reference here. The work is being prosecuted so rapidly, especially in Cambridge and Berkeley, that new results are appearing almost weekly. We may expect to hear soon of new classes of projectile, to add to the proton, diplon and helium ion. There seems no reason why lithium and all the other ions used by Aston should not be produced in far greater quantities and accelerated by far higher potentials than those needed in mass ray analysis. The chemistry of nuclear reactions will soon become as familiar a field as any other kind of inorganic chemistry.

§ 7. THE α-RAY SPECTRUM AND NUCLEAR STRUCTURE

Until recently it was stated that the α-particle from a particular parent element had a definite velocity, or energy, characteristic of that element, or, in other words, that the α-ray spectrum consisted of a single line. During the last few years it has been proved, firstly, that, in addition to the well-known particles of characteristic range, which will be called the main particles, there are a comparatively few discrete particles with much longer ranges, and secondly that the main particles have not all one range, but are distributed among a few close ranges. We speak of long-range particles and of the fine structure of the α-rays. It must be borne in mind that the long-range particles are very rare in comparison to those that make up any line of the fine structure of the main line.

The long-range particles have been investigated by Rutherford and his collaborators, F. A. B. Ward, C. E. Wynn-Williams, W. B. Lewis, and B. V. Bowden[85,86,87,88,89]. The particles are detected by means of the ionization produced in a small ionization chamber, the charge of potential consequent on the charge released being amplified by the methods referred to in § 2. The difficulty is to count in the presence of strong γ-radiation. The differential ionization chamber of Rutherford, Ward and Wynn-Williams, described in § 2, has rendered great service in overcoming this source of trouble. The experiment is made by gradually changing the amount of absorbing material between the source and the chamber, and measuring the variation with absorption of the number of kicks per unit time. In the earlier experiments the amplified current was made to register on a photographic film by means of a special oscillograph; later it was made to operate a mechanical counter, by means of a thyratron device. It is easy from the plotted values to determine the ranges of the particles. The long-range α-particles from certain elements, notably Ra C and Th C, have been analysed by Nimmo and Feather[90], and by Philipp and Donat[91], by another method, namely the Wilson chamber.

The fine structure of α-ray groups was first measured by Rosenblum[92,93,94,95] in the case of the transition Th C to Th C″. He used the great magnet at Bellevue near Paris, which weighs 120 tons and uses a current of 400 amperes at 240 volts—the very respectable power of 96 kw. With a 1 cm. gap and pole-pieces 25 cm. in diameter a field of 36,000 gauss, uniform to within 1 part in 500, can be produced over an annular space 2 or 3 cm. wide, and 20 cm. in diameter: with pole-pieces 75 cm. in diameter, and a gap of 5 cm. or so, a field of 24,000 gauss can be produced which is sufficiently constant over a space 35 cm. in diameter. It is less the intensity of field than its effective size that is impressive, and that is essential for the measurements in question. Rosenblum used a photographic plate in his earlier experiments: recently he has adopted a Geiger counter, which furnishes more precise results as to the relative intensity of the different classes of α-rays.

A magnet of special type which is in use in the Cavendish Laboratory[96] also allows the fine structure of α-rays to be measured with great accuracy. It is designed to produce not the usual cylindrical field but an annular field only, this being what is required for the measurements in question. The cross-section is shown in figure 8 (*a*), the whole magnet being produced by a rotation of the figure about a vertical central axis. The coils *EE* are wound round the central core *C*. The annular field produced in the gap *GG* is about 40 cm. in radius and 5 cm. in width. If a source is introduced into the field through the hole *H* and a detector be arranged diametrically opposite to it, through the hole *J*, then, by adjusting the field, α-rays of any given velocity can be brought into focus at the detector, as shown in figure 8 (*b*). The focussing at opposite ends of a diameter when the rays are bent into a semicircle in a magnetic field has long been used for β-rays, and was first applied to α-rays by Rosenblum. The beam is limited by diaphragms so as to lie entirely within the uniform part of the field. The detector is an ionization chamber, connected to an amplifier and a thyratron device operating a mechanical counter. The field is varied, and the number of particles registered observed for different

fields. The maxima enable the velocities of the different groups to be determined. Two homogeneous groups differing in velocity by only 1 in 5000 can be resolved. A high vacuum is, of course, maintained in the gap in all experiments of this type.

The long-range α-particles are very few in comparison with the main particles

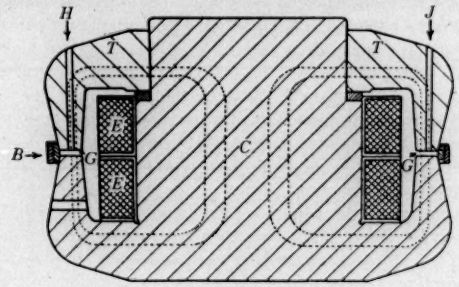
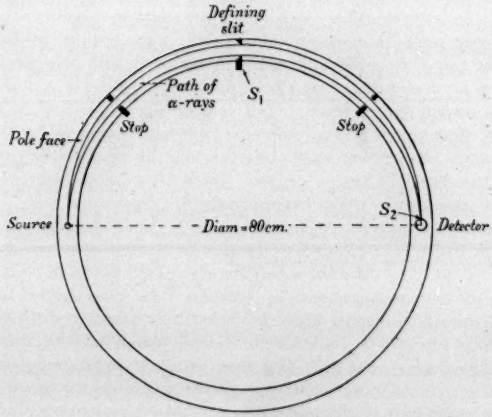


Figure 8 (*a*).



[*Reproduced by courtesy of the Royal Society.*

Figure 8 (*a*).

as has been already mentioned. We may take Ra C as an example. Besides the well-known particles of range 6·88 cm. (in air at 15° and 760 mm.) there are various groups, some nine in number, with energies between 8·30 and $10{\cdot}62 \times 10^6$ ε volts. The ranges vary from 7·79 to 11·52 cm. By the magnetic deflection the velocity and hence the energy, which is what is required for the consideration of nuclear

levels, is found directly, which gives greater accuracy than a deduction from the range. There are about 17 of the most plentiful of the long-range particles, those of range 9·04 cm., to every million of the main particles: the other long-range particles occur in the proportion of 1 or less to a million of the main particles.

Before the bearing of these recent α-rays on nuclear structure is considered it is advisable to discuss briefly the origin of the β-ray line spectrum. This is a matter which has been studied in great detail by C. D. Ellis[97, 98], who has in particular compared the natural β-ray line spectrum of each radioactive element with that excited by the γ-rays from that element in other heavy elements on which they fall—the so-called secondary β-ray spectrum. In many cases the natural lines coincide with those excited in different metals, such as platinum, when allowance has been made for the energy necessary to liberate the electron from its particular binding, say in the K level, of the atom. This can be directly explained by supposing that in both natural and secondary spectra the electrons are photoelectrically released by certain γ-rays, of frequency ν_1; $\nu_2 \ldots$, so that their observed energy is

$$h\nu_1 - K,\ h\nu_1 - L, \ldots;\ h\nu_2 - K,\ h\nu_2 - L, \ldots;$$

$K, L \ldots$ being the energies of binding in the corresponding levels, which vary, of course, from element to element. There are, however, as first pointed out by Thibaud[99], certain lines in the natural spectrum to which no corresponding lines can be found in the secondary spectrum, which would not be the case if the primary β-particles were released by γ-rays, and further there are certain difficulties in the way of supposing that γ-rays can act photoelectrically on electrons in the atom in which they originate. Ellis has therefore attributed the release of the primary β-particles to a process which he calls "internal conversion." By this is meant that a quantum energy charge in the nucleus of a given atom may lead directly to the expulsion of an electron from the same atom with this energy, allowance being made for the energy of release, as above.

The internal conversion is supposed to take place owing to the fact that the electron system is coupled to the nucleus, so that the atom is to be treated as a whole, governed by the quantum rules of wave mechanics. The nuclear state is characterized by certain energy levels, to each of which an azimuthal quantum number i is attributed, giving the spin of the nucleus in that state. When the nucleus passes from one of these energy states to another, without any corresponding change in the electron system, a γ-ray is emitted, whose $h\nu$ equals the change of energy. On the other hand, the change of energy of the nucleus may be transferred directly to an electron (internal conversion). The nuclear transitions are governed by certain selection rules, just as are the changes in the azimuthal energy states of the electron system when optical lines are in question. These selection rules are different according as the transition is solely concerned with the nucleus or as it involves the electrons. Thus a transition $0 \to 0$, $\Delta i = 0$, is a very weak, or improbable transition as a nuclear change alone, so that a γ-ray corresponding to it will be missing, or very feeble, but when, owing to the coupling, it is considered as involving the electron energies, say of a K level, it is strong. It is in this way that

the occurrence of β-ray groups in the natural spectrum, to which there are no corresponding groups in the secondary, or excited, β-ray spectrum, is explained.

In this connexion a quantity known as the coefficient of internal conversion plays an important part. Suppose for simplicity that we have a number N of nuclei, of given atomic species, in each of which one quantum transition takes place. A certain number p of γ-rays, of energy $h\nu$, will be emitted, and a certain number of electrons q will be liberated by internal conversion, where $p+q=N$. The γ-rays can be detected by the liberation of secondary electrons in foreign atoms, as explained, and their number thus found. On this basis C. D. Ellis[100] has devised methods for measuring the ratio $q/p=\alpha$, which is defined as the coefficient of internal conversion*.

Returning to the selection rules, we find that transitions for which $\Delta i = \pm 1$ can occur, and for each of these the internal conversion coefficient α can be found. Such transitions are called dipole transitions, for reasons that need not detain us here. In contradistinction to optical spectra (where $\Delta - = \pm 2$ is forbidden) transitions for which $\Delta i = \pm 2$, called quadrupole transitions, can occur for the nucleus†. If the measured $\alpha = q/p$ be plotted against $h\nu$ for all transitions experimentally observed they are found to lie on two separate lines, one corresponding to dipole transitions and the other to quadrupole transitions. These two lines are important in sorting out levels, as they furnish criteria in attributing values of i to the levels, since the Δi's must conform to these curves.

Rutherford and Ellis[101] have pointed out a connexion between nuclear levels, as established by α-ray measurements, and the γ-ray spectrum, as revealed by Ellis's measurements on β-ray line spectra. It is assumed that the nucleus can exist in various excited states, as has been explained, but for the moment we will consider only one excited state, and the ground state of the parent atom A and of the product atom B, as indicated in figure 9 (*a*). Nucleus A can pass from the state represented by a, to the ground state a_0, the energy being transformed into a γ-ray of frequency ν, or it can emit an α-ray while in state a, and pass to the ground state b_0 of the product atom. Such an α-particle will clearly have more energy than one emitted from A while in the ground state a_0, and is a "long-range" α, of infrequent occurrence. Ordinarily the atom A emits an α-particle when it is in its ground state, such α's being what we have called the main particles. The atom B may be formed in its ground state b_0 or in an excited state b, the corresponding energy differences give the "fine structure" of the α-ray spectrum, which in the simple case represented will consist of two lines only. Levels have now been assigned on this basis which account in a satisfactory quantitative way for all the long-range α's and γ-rays.

There is no fundamental difference between the long-range α's and the fine-structure α's, all of which are given by passage from a stationary energy state of a

* The use of the symbol α is not very happy, in view of the constant occurrence of α-rays in these discussions, but it is always used by the original writers, and so is retained here.

† The theory of this has been worked out by Mott and his collaborators. For full references see J. B. Fisk, *Proc. roy. Soc.* **143**, 674 (1934).

nucleus to a stationary energy state of the nucleus formed by loss of an α-particle. The difference of character is largely due to the fact that the nucleus has a very short life in the excited state, as can be easily seen by consideration of the experimental results. The atom in state a can either pass to a_0, giving out a γ-ray, or it can emit a long-range α-particle. The probabilities of γ-ray and α-ray transition in time dt can be denoted by $g.dt$ and $a.dt$ respectively, so that if G and A are the average numbers of γ and α units emitted by a given number of nuclei in unit time

$$\frac{g}{a}=\frac{G}{A},$$

which can be observed experimentally, and is large, of the order 10^5 or 10^6. To fix the probability of an α-ray disintegration occurring in a nucleus in the excited state, of known excess of energy E' over the ground state of the product nucleus, we can make use of Gamow and Houterman's[102] improved formula of Geiger-Nuttall type, connecting disintegration constant and energy of expelled particle. The energy

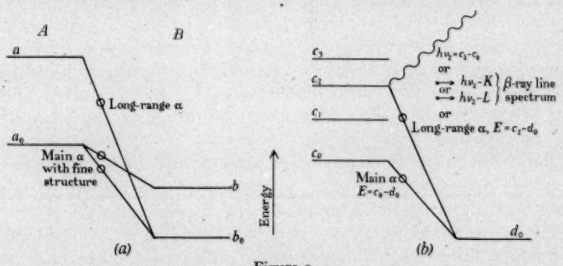


Figure 9.

E of the long-range particle being large, the probability of an α-ray disintegration is large. For example, with a particle of range 7·79 cm. the α-particle leak, which expresses the probability of a transition from the level in question, is about 10^8: for a particle of range 10·21 it is 3×10^{11}*. The probability of a γ-ray transition in 1 sec. is, then, of the order of 10^{13} to 10^{17}, which can be alternatively expressed by saying that the life of the excited nucleus is very short, of the order of 10^{-13} sec. or less. This is to be expected on general grounds, since the life of an excited atom, as evidenced by consideration of the optical spectra, where the frequencies are much lower, is of the order 10^{-8} sec. Certain general arguments go to show that the life should decrease rapidly with rise of frequency.

We may now consider an actual case of the nuclear structure of a radioactive element and its product, that of Ra C and Ra C′. A Ra C nucleus which has emitted a β-particle becomes a Ra C′ nucleus, which in its turn emits an α-particle and becomes Ra D, the half-life period of the Ra C′ nucleus being of the order 10^{-6} sec.

* Cf. the familiar case of the α-particle of range 8·60 cm. from Th C′, the half-life period of which is given as of the order 10^{-11} sec. in the text-book. This is based on the older Geiger-Nuttall formula.

There is also an emission of γ-rays, accompanied by a β-ray line spectrum, analysis of which enables the lines to be attributed to internal conversion. We have further the long-range α-particles observed by Rutherford, Ward and Lewis. All these facts can find quantitative expression in the following scheme.

The Ra C′ nucleus is found in one of several possible excited states, indicated (not to any scale) in figure 9 (*b*) by $c_1, c_2, c_3 \ldots, c_0$ being the ground state. The Ra D nucleus also probably has possible excited states, but none have yet been measured, and so the ground state d_0 alone is shown. Now when the Ra C′ nucleus passes from an excited state whose energy level is c_2, say, to the ground state we have various possibilities. A γ-ray of energy $h\nu_2 = c_2 - c_0$ may be emitted. Or, by internal conversion, an electron may be emitted from one of the electronic (X-ray) levels of the Ra C′ atom, with energy $h\nu_2 - K$ or $h\nu_2 - L$, or ...; K, L ... being the energy required to free the electron from the level in question. Or an α-particle may be emitted from the Ra C′ atom in the excited state, in which case the nucleus will become a Ra D nucleus (in the ground state) so that the whole energy c_2 is carried off by the α-particle, which is a so-called long-range α-particle. Experiment shows, however, that this is an infrequent occurrence. In general the Ra C′ nucleus reverts to its ground state with emission of a γ-ray, and the α-particle is discharged from the nucleus in the ground state, which becomes a Ra D nucleus in the ground state, as shown. All the experimentally determined energy changes and intensities agree well with this scheme.

The fine structure of the α-rays, not observed with Ra C′, occurs when the nucleus which results from an α-ray transformation can be formed in various excited states. For instance, Th C gives out an α-ray with range about 4·7 cm., for which a fine structure has been established by Rosenblum and by Rutherford and his collaborators. This is due to various possible excited states of the Th C″ nucleus. If the Ra D nucleus resulting from Ra C′ were formed in an excited state in an appreciable fraction of the main α-ray transformations, there would be various levels between c_0 and d_0, transitions to which from c_0 would give a fine structure.

In general the ordering of nuclear levels is a complicated matter calling for an elaborate comparison of a number of different lines of evidence. If the only criterion to be satisfied was that the energy differences should agree with emitted β- and γ-ray energies, the level schemes would be very uncertain, as the percentage precision of the measurements is not very high when it comes to comparison with differences. Besides energy conditions, however, in which we include the energies of the long-range and fine-structure α-rays, connecting preceding and succeeding disintegrations, the levels have to be ordered so as to satisfy the measured intensity relationships, both for the β-rays and the γ-rays. Further, there is the question of spin, and the related question of the ordering into dipole and quadrupole transitions, so as to satisfy the internal conversion conditions, to which reference has been made. Again, Fermi has calculated certain interrelations between the spin values of parent and product nuclei. Making use of all these guides, success has been attained in elaborating nuclear schemes for intricate cases, such as that of the alternative Th C to Th D transmutations[103, 104].

§ 8. ANALYSIS OF THE EXPELLED PROTONS

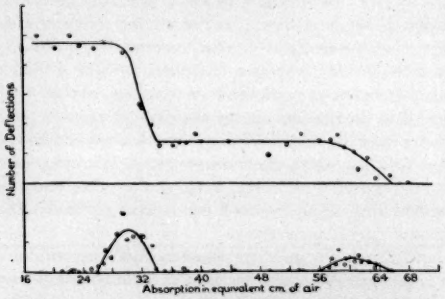
In the case of certain elements, especially aluminium, an extensive study has been made of the protons expelled by α-ray bombardment. The importance of knowing such things as the effectiveness of α-particles of different energy in dislodging protons, and the energies of the protons, is that this type of information throws light on the different possible energy states of the light nuclei. The knowledge afforded is crude compared with that which we have of the more important radioactive nuclei, where the intensities of the radiations make measurement relatively easy, but at the present time the method is the only one available.

There are various ways in which the experiments can be conducted. Factors which can be varied are the thickness of the target and the energy of the bombarding α-particles. What is measured is the number of protons having a given range, or energy—the so-called distribution in range. With a thin target we have only α-particles of the incident energy to consider, but with a thick target many of these are slowed down before they effect a transmutation, so that we have α-particles of all energies less than the incident energy also present. In addition to the number of protons, then, we have the three following variables: range, or penetrating power of protons; range of α-particles; thickness of target. We can plot the number of protons against any one of these three, keeping the other two constant for any particular curve, but varying one or both from curve to curve. Different experimenters have adopted different methods.

Rutherford and Chadwick, in their pioneer work on artificial disintegration, called attention to the long-range protons emitted from aluminium under α-particle bombardment. Bothe and Fränz[105] were the first to show that the proton radiation could be divided into groups, and during the last few years these groups have been the object of a large number of researches. All the experimenters except Haxel seem to have used the α-particles from polonium, which has the advantage that it emits no β-rays, but the disadvantage that the energy of the α-particle is relatively low—range 3·9 cm. One of the first workers to follow Bothe and Fränz was Pose[106, 107, 108, 109], who, with α-particles of various energies (ranges 3·9 downwards), measured the distribution in range of the expelled protons from sheets of aluminium sufficiently thick to absorb all the α-particles, which means that in any one experiment the protons were liberated not only by α-particles of the incident energy, but also by those which had penetrated some distance, and so were of less energy. He also carried out experiments with different thicknesses of foil.

Chadwick and Constable[110], following the earlier work of Chadwick, Constable and Pollard[111], have in particular used very thin sheets of aluminium (0·8 mm. air equivalent) to confirm the results of their experiments on thicker sheets, the thin sheets giving results for incident α-particles which are practically monoenergetic. Figure 10, which illustrates their results with α-particles of range 3·9 cm., is given as one of the simplest distribution-in-range curves, which are characteristic of practically all work in this field. In the upper part of the diagram is plotted the

number of deflections given by the registering device of the counter against the air equivalent of the absorbing screens inserted between the source of protons (aluminum foil) and the counter. All particles of a fixed range clearly register until the thickness of absorbing screens becomes about equal to their range, when a small increase of thickness makes the number registered diminish rapidly. The horizontal line means that the particles are being transmitted in fixed number; the sudden drop means that the absorbing thickness corresponds to the range of the particles. The upper curve shows, then, that there are two groups of particles, of maximum ranges 34 and 66 cm. respectively. The lower curve gives the result of a differential count. In this method only large deflections, corresponding to heavy ionization, are reckoned, and heavy ionization only takes place at the end of the path of a particle. As will be seen, the two methods agree exactly.



[*Reproduced by courtesy of the Royal Society.*

Figure 10.

Pose deduced from his measurements that there were three groups of protons to be detected from aluminium, a group of short range, about 26 cm., whose range and numbers vary continuously with the energy of the primary α-particles, and two groups of fixed range, 47 and 58 cm., which are only excited by α-rays within certain narrow limits of energy, the most effective α-rays being those of about 2·2 and 3·2 cm. range, respectively. Chadwick and Constable also find that α-particles of certain energies are particularly effective in expelling protons, but their results do not agree in any detail with those of Pose, which other observers, for instance Steudel[112] and M. de Broglie and Leprince-Ringuet[113], have also been unable to confirm. According to Chadwick and Constable there are four groups of effective α-rays, each one of which is able to expel two groups of protons, the range of the protons in these groups being the less the lower the energy of the activating α-particles. In addition, there is the continuous group of short range.

Haxel[114, 115] has carried the matter farther by using α-rays from Ra C and Th C, thus extending the range up to 8·6 cm. He avoids the difficulties attaching to counting protons in the presence of an overwhelming number of penetrating

β-particles by using a Geiger-Müller tube counter with a linear amplification of the ionization, which Geiger and Klemperer[116] had shown to be possible over a certain range. The ionization produced by the β-particle is very small compared to that of the proton, which therefore can be detected as such. He found that the α-particles of high energy E_α produce four ranges of protons, while those of smaller energy produce only two, as found by Chadwick and Constable. The energy E_p of the protons produced varies linearly with E_α. There are thus four parallel lines when E_p is plotted against E_α, and Chadwick and Constable's values lie exactly on the upper two, as shown (figure 11).

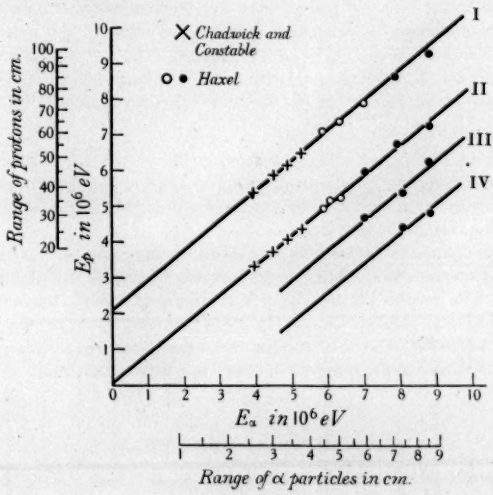


Figure 11.

These four groups cannot be explained on the hypothesis of two proton energy levels in the aluminium nucleus, and two in the silicon nucleus formed by the capture of an α-particle and the loss of a proton, for if that were so the energies E_p at constant E_α would be in two pairs with a constant separation, which they are not. They must therefore be due to four levels in one, and one level in the other, of the two nuclei $^{27}_{13}$Al and $^{30}_{14}$Si. Now Bothe and Becker found that aluminium emits γ-rays as well as protons, to explain which we must assume the formation of an excited silicon nucleus, emitting a γ-ray when it returns to its ground state. It appears, therefore, that the results are most easily explained by four levels in the silicon nucleus, the differences of which are given by figure 11.

Haxel's experiments to date do not throw any light on the question of preferential excitation by α-particles of particular energies, since he has measured only

the range of the protons, and not their numbers, but the experiments of Pose and of Chadwick and Constable seem quite definite on this point, although there is a lack of agreement on other matters. This effect, as pointed out by Gurney[117], can be explained on Gamow's theory, as a kind of resonance between the wave system of the α-particle and an energy state of the nucleus. It has been pointed out that the transmission of a particle through the potential barrier has some analogy to the transmission of light through a single reflecting surface, but if a particle approaches the nucleus it encounters, effectively, two barriers close to one another, the opposite sides of the crater. Just as in optics the placing of a second reflecting surface close to the first, as in the Fabry-Perot arrangement, causes preferential transmission of certain wave-lengths, so the space between the boundaries may govern a preferential admission of certain energies. The so-called resonance levels in the nucleus can, then, be theoretically justified. If the energy of the α-particle when captured is E_a^0, and that of the proton before dislodgment is E_p^0, then

$$E_p = E_a + E_a^0 - E_p^0,$$

which for fixed levels gives the linear relationship between E_p and E_α found by Haxel. The continuous group of protons of short range for boron has been tentatively explained as due to dislodgment of a proton by the α-particle, without capture. This is a little uncertain.

Protons of definite range, without continuous background, have been established in the case of several other light elements by Pose, Chadwick and his collaborators, and others. The results are mostly of a preliminary nature, but are extremely promising, and seem to give hope that the level schemes and mechanism of expulsion of particles will soon be generalized for the non-radioactive atoms, and possibly linked up with those of the radioactive atoms in a general scheme.

§ 9. THE SPIN OF THE NUCLEUS

In a number of spectral lines, especially of the heavy elements mercury, thallium, lead and bismuth, a fine structure has been established which, to distinguish it from the ordinary multiplet structure, is called the hyperfine structure. The heavy atoms are advantageous because not only is the hyperfine structure more open, but also, on account of the low velocity and diminished Doppler effect, the component lines tend to be very narrow. The frequency separation of the components is very small, the order being only about 1/80,000 of the separations in the ordinary multiplet structures. The wave-number interval between the extreme members is, for instance, 2·35 cm.$^{-1}$ in the case of the bismuth line λ 3596, $2F'_{\frac{1}{2}} - 4d''$, the structure of which is, with that of the bismuth line λ 4722, $^2D_{\frac{3}{2}} - 2s$, shown diagrammatically as an example in figure 12 (*a*). The frequencies of the components can be expressed as differences of term frequencies by a level scheme, as exemplified in the analysis of the two lines of figure 12 (*a*) given in figure 12 (*b*), and, of course, it is this splitting of the levels that is the fundamental thing.

This splitting can be explained, as was first suggested by Pauli[118], by considering

the effect, on the nucleus, of the magnetic field produced by the circulation of the outer electrons. The general outline of the theory of the effect is as follows. Suppose that the nucleus has a mechanical moment of momentum of quantum number I, and a magnetic moment μ. The extra-nuclear electron swarm will produce a magnetic field H at the nucleus due partly to the circulation of the electrons

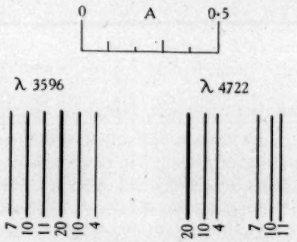


Figure 12 (*a*). Hyperfine structure of two bismuth lines.

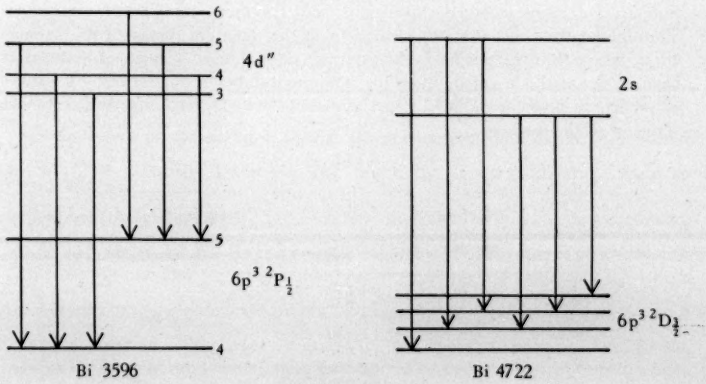


Figure 12 (*b*).

round the nucleus, and partly to the intrinsic magnetic moment of each electron. The electron swarm likewise has a moment of momentum characterized by the quantum number J. The nucleus is coupled to the electron swarm in such a way that the total moment of momentum F of the system is given by the vector composition of J and I, only resultants differing from one another by a unit being admitted, in the usual way. Thus if $I > J$ there will be $2J+1$ possible values of the resultant moment, or if $J > I$ there will be $2I+1$ values. Corresponding to

each there will be a different energy state of the system, so that each level will be split into $2J+1$ or $2I+1$ terms. Simple considerations of angle show that the difference of energy ΔW of the system of moment F from the undisturbed value is given by

$$\Delta W = \mu H \frac{F(F+1) - (I+1) - J(J+1)}{2IJ}.$$

J is, of course, different for different terms, and must be determined from spectral considerations. There is an F selection rule for the transitions, analogous to the ordinary spectral selection rules: it is $\Delta F = 0$ or ± 1 with the special transition $0 \to 0$ forbidden.

We can apparently find I from the number of components if we are sure that $J > I$, but actually there are certain difficulties which space will not allow us to discuss. The general method of finding I in practice is from the relative intensities of the different components, which depend upon ΔJ and ΔF. The relationship is, for instance, particularly simple when we have a transition between a term which remains simple and a term split into two levels, when the ratio of the intensities of the two lines is given by

$$[2(I+J)+1]/[2(I+J)-1].$$

Similar, but more involved expressions hold for more complicated transitions. Throughout the considerations remind us of the quantum theory of the Zeeman effect, where the magnetic field is an external one. Schüler[119] and his collaborators have, by the methods stated, found the quantum number of nuclear spin for a series of nuclei, as given in table 1. Such moments are of great importance for the statistics of the nucleus.

Table 1.

Element	Nuclear spin number	$\frac{\mu}{\mu_0}$ from	
		Fermi	Goudsmit
^{7}Li	3/2	+1/575	
^{23}Na	3/2	+1/880	
^{27}Al	1/2		+1/875
^{63}Cu	3/2	+1/780	
^{65}Cu	3/2	+1/780	
^{69}Ga	3/2	+1/860 } 1/765	
^{71}Ga	3/2	+1/670 }	
^{87}Rb	3/2	+1/660	
^{121}Sb	5/2		+1/680 } 1/775
^{123}Sb	7/2		+1/875 }
^{133}Cs	7/2	+1/700	
^{209}Bi	9/2	+1/520	
^{85}Rb	5/2	+1/1350	
^{203}Tl	1/2	+1/1300	
^{205}Tl	1/2	+1/1300	
^{115}In	9/2	+1/350	
^{197}Au	3/2 ?	+(1/1010 ?)	
^{75}As	3/2		+1/2040

Table 2.

Element	Nuclear spin number	$\frac{\mu}{\mu_0}$ (Fermi)
^{87}Sr	$\geqq 3/2$	$-1/1880$
^{117}Sn	$1/2$	$-1/2030$
^{119}Sn	$1/2$	$-1/2030$
^{137}Ba	$3/2$	$+1/1950$
^{111}Cd	$1/2$	$-1/3500$
^{113}Cd	$1/2$	$-1/3500$
^{199}Hg	$1/2$	$+1/4000$
^{201}Hg	$3/2$	$-1/3600$
^{207}Pb	$1/2$	$+1/3500$

As regards the magnetic moment, we cannot deduce it unless we know H, the field produced at the nucleus by the electron swarm. During the past year this has been satisfactorily worked out on wave mechanical principles by E. Fermi and E. Segrè[120], whose theory satisfies the criterion that the consideration of different spectral terms of one atom shall give the same magnetic moment for the nucleus of that atom. A similar calculation of nuclear magnetic moments was carried out independently by Goldsmit[121], and there is a satisfactory agreement between the moments derived by these two investigators.

It is to be expected that the magnetic moment of the nucleus will be very small. The value of the Bohr magneton is $\mu_B = \frac{1}{2}\frac{e}{m}\frac{h}{2\pi}$, and the mass of the proton is 1840 times that of the electron, so that the proton magneton might be expected to be $\frac{\mu_B}{1840}$. Actually Estermann and Stern, by an extraordinarily ingenious experiment, have shown it to be 2·5 times this value, or $\frac{\mu_B}{736}$, a very surprising result, so far unexplained. As can be seen from table 1, in the case of elements with odd mass number and odd atomic number the experimental value of μ/μ_B, which are not very exact, approximates in most cases to either $1/736$ or $\frac{1}{2}\cdot\frac{1}{736}$, as pointed out by Schüler and Westmeyer. ^{115}In, ^{197}Au, ^{75}As appear to be exceptions. In the case of elements of even atomic number and odd mass number μ/μ_B approximates to $1/1840$ or $\frac{1}{2}\cdot\frac{1}{1840}$. Whether this is a valid generalization remains to be seen.

The study of the hyperfine structure is complicated by an isotope effect, since different isotopes of the same element may have different nuclear moments. This isotope effect was first established by Schüler and others[122, 123]. Schüler and his collaborators have devoted much attention to this point, and Schüler and Jones[124] succeeded in establishing from the hyperfine structure the existence of a new lead isotope ^{204}Pb, which was later confirmed by Aston. This is striking evidence of the reliability of the interpretation of this complicated phenomenon.

There is a second method for deriving the moment of momentum of the nucleus, namely from band spectra[125, 126]. It is well known that in the spectra of molecules made up of two like atoms the lines in the electronic bands alternate in intensity: in certain cases, notably He_2 and O_2, the alternate lines are missing altogether. Now it is well known that the wave function of a state of a molecule is a product of three parts: one pertaining to the electrons; a second depending upon the separation of the two nuclei; and a third involving the angles which fix the line joining the two nuclei, and so representing the rotational possibilities of the molecule as a whole. If the nuclei are treated merely as mass points the wave equation must be symmetrical with respect to their coordinates, and the wave function representing the stationary state of such a molecule is either symmetrical or anti-symmetrical, i.e. when the nuclei are interchanged it either remains unaffected or changes sign. The rotational states being the ones involved in the lines in question, the quantum number J of the total moment of momentum is the one concerned, and there will be $(2J+1)$ states, giving the statistical weight of the level, if we take no account of the symmetrical and anti-symmetrical possibilities. When we take these into account we can express the intensities by calling the statistical weights $g_s(2J+1)$ and $g_a(2J+1)$ respectively for the cases where the state is symmetrical and anti-symmetrical in the coordinates of the nuclei.

The results of line spectroscopy teach us that only one kind of wave function, the symmetrical or the anti-symmetrical, can occur, so that if the nuclei are treated as points, fully fixed by space coordinates alone, either g_s or g_a is zero, and alternate lines are completely missing. Apparently, then, the nuclei of He and O can be treated as points and have no spin. If, however, we attribute spin to the nuclei, we have a new wave function for the molecule which can be symmetrical or anti-symmetrical, and so multiplying by this function can render the anti-symmetrical point-coordinate function symmetrical, and *vice versa*, when all coordinates are considered. If, according to Bose-Einstein statistics, we admit only states symmetrical in all coordinates of the nucleus,

$$\frac{g_s}{g_a}=\frac{I+1}{I},$$

while if, according to Fermi-Dirac we admit only states anti-symmetrical in all coordinates, we have

$$\frac{g_s}{g_a}=\frac{I}{I+1},$$

where I is the quantum number of the nuclear spin. In either case the ratio of the intensity of faint to strong lines is the same, and allows us to find I. To know which of the quantities g_s or g_a is the bigger, i.e. if the symmetrical or anti-symmetrical states have the greater weight, we must know the wave function.

Fortunately the question as to the type of statistics can also be decided from the Raman effect of the gas, as was pointed out by Heitler and Herzberg[127], and, e.g., it was from measurements of this kind that the statistics of the nitrogen nucleus was

decided. This is of importance in connexion with the questions discussed in the last section of this report.

From what has been said it is clear that the measurements of the relative intensity of alternating lines gives the value of I. Thus for nitrogen the ratio as found by Ornstein and van Wijk[128, 129] is 2 : 4, which gives $I=1$.

The band spectra show another important effect bearing on the nuclei, namely the isotopic effect. Both for vibrations of the two nuclei and for rotations of the molecule as a whole the nuclear masses are of importance, supposing that the internuclear force and distance have, respectively, a fixed value, as they will have with two isotopes, since they depend essentially upon the net nuclear charge Z. As regards the vibrational effect it can easily be shown that the fundamental frequency ν_0 is inversely proportional to $\sqrt{M}$, while, from wave mechanics, the energy levels are given by

$$E_{\text{vib}}=h\nu_0\,(n+\tfrac{1}{2}), \qquad n=0,\ 1,\ 2\ \ldots.$$

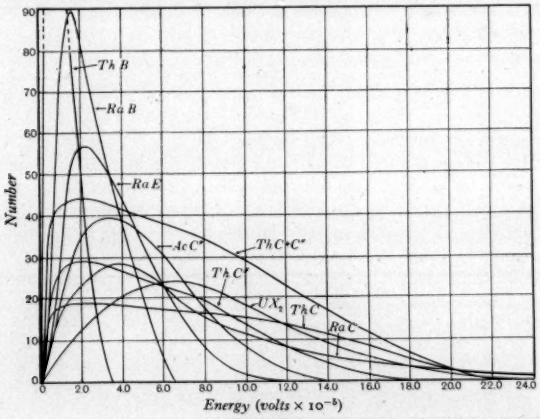
The vibrational levels are closer for the heavier isotope than for the light one. As regards the rotation, the moment of inertia being proportional to M the fundamental frequency is inversely proportional to M. The rotational levels are also closer for the heavier isotope. Actually the calculations are rendered complicated by the action of the electronic movements on the vibration and the rotation of the molecule as a whole, but the isotope effect has been worked out in great detail, and the structures observed have been satisfactorily accounted for.

The isotopes and their masses are determined by the numbers and position of the satellites caused by their presence: an estimate of their relative abundance can be made from intensities. The band spectra have proved particularly serviceable in the matter of the new isotopes of oxygen.

§ 10. THE β-RAY SPECTRUM. THE NEUTRINO

In the case of α disintegration the particles issue from the nucleus with a few discrete energies, one particular energy being commoner than the others. We have, as it is said, a line spectrum for the energy of the α-particles, which, as we have seen, is explicable in terms of a few energy levels within the nucleus. When we turn to β disintegrations the situation is less satisfactory. In the case of all substances which disintegrate with emissions of a nuclear electron the electrons do not possess a few discrete energies but, within the limits of resolutions possible to-day, furnish a continuous spectrum when n is plotted against energy E, $n.dE$ being the number of particles possessing energy between E and $E+dE$. There is, of course, in the case of nearly all bodies, a line spectrum of β-rays, amounting in all to a few per cent of the total emission, imposed on this continuous background, but this does not invalidate the statement that all energies within a certain range are represented. The facts are, then, clearly not to be explained on the ground of a few nuclear levels, and, as will be seen, all solutions that have hitherto presented themselves offer difficulties.

In the first place we cannot assume that each nucleus emits several electrons of different energy, or, to take a simple possibility, that each nucleus emits two electrons which share the available energy between them in a random ratio, which varies from atom to atom, for careful experiment has shown that only one electron is given out per atom. In any case the emission of more than one nuclear electron would completely upset our ideas on the periodic table, which have been so successful. We must, then, restrict ourselves to one electron for disintegration. Now the energy distribution, given for several elements in figure 13, is characterized, as shown in particular by Sergant[130], by an upper limit to the energy in the case of each element. It seems possible that all the electrons might leave the nucleus with the same energy, but that part of this energy might be absorbed in the atom itself



[*Reproduced by courtesy of the Cambridge Philosophical Society.*]

Figure 13.

or in the surrounding atom, the amount differing from atom to atom. On this hypothesis the initial energy would clearly have to be that corresponding to the upper limit. The view can be tested by measuring the heating effect, the specimen being held in a thick-walled vessel, to absorb any secondary radiations set free in the absorption process. If the β-radiation is initially homogeneous the heating should correspond to the maximum energy per atom; if the measured β-ray energy represents the energy with which the β-particles are discharged from the nucleus the heating effect would represent the average energy, $\frac{\int ndE}{\int dn}$ given by the distribution curve, per atom. There is a considerable difference between the two cases, and the measurements of Ellis are quite decisive in favour of the latter alternative. He used a Ra E preparation in a lead calorimeter with walls 1 mm. thick, since the transformation Ra E to Ra F is a β-ray transformation unaccompanied by

complications such as γ-rays. Ellis's results have been confirmed by Meitner and Orthmann[131]. The last possibility, that there might be differences among the nuclei with respect to either the initial or the final state, must also be ruled out. Differences in the nuclear structure would manifest themselves in other ways, as, e.g., in difference of ranges of the α-particles from Ra F (polonium), which are not found.

If, then, the β-particle emission is the only process that takes place, or if there is no other form of energy emitted than that which can be absorbed by the lead walls of the calorimeter in Ellis's experiment, the conservation of energy cannot hold for β-ray disintegration. A way of avoiding this unpleasant conclusion is to suppose that there is some form of radiation, extremely difficult to detect, which carries off the surplus energy representing the difference between the maximum and the average energy of a β-particle. Another difficulty, however, concerns the angular momentum of the nucleus. Experiment shows that nuclei of even mass number have an even quantum spin number I, while nuclei of odd mass number have an odd value of I. When the β disintegration takes place the mass number does not change, so that the spin difference between parent substance and product nucleus is 0 or an even multiple of $\frac{1}{2}\frac{h}{2\pi}$. On the other hand, the spin of the electron is $\frac{1}{2}\frac{h}{2\pi}$, while any angular momentum in an orbit is a multiple of $\frac{h}{2\pi}$ and so an even multiple of $\frac{1}{2}\frac{h}{2\pi}$. There is, therefore, a discrepancy spin of $\frac{1}{2}\frac{h}{2\pi}$ in the β disintegration.

Pauli suggested that we can get round the difficulty by postulating a particle endowed with energy, momentum, and angular momentum (spin) which carries off the surplus energy and the angular momentum of $\frac{1}{2}\frac{h}{2\pi}$. The name neutrino has been given to such a hypothetical particle. From complicated considerations of the β-particle emission, assumed to be the consequence of a neutron-proton transition, Fermi concludes that the mass of the neutrino must be much smaller than that of the electron, perhaps zero. The neutrino has no charge, and the ionization produced must be extremely small, not exceeding one pair of ions per fifteen kilometre length of path. This, and the lack of interaction with nuclei or other particles, renders the detection of the particles extremely improbable, if not impossible.

A particle which, like the hypothetical neutrino, is *a priori* unlikely to be ever detected is not a very sympathetic entity in these days. In fact it may be said—no possibility of detection, no real existence. Nevertheless the neutrino has its uses, as it serves as a convenient summary of the conservation difficulties offered by β-ray emission. It is, as it were, a bag in which we can put all the troubles together, although it may have no more material existence than the troubles. Whether we say that it accounts for the failure of the conservation laws, or is just a way of expressing the failure of these laws, is a matter of taste. If we accept the hypothesis we have to say that when our heavy particle passes from the quantum state neutron to the quantum state proton, an electron and a neutrino are emitted.

§ 11. ARTIFICIAL RADIOACTIVITY

A new chapter in the history of the nucleus has been opened by the discovery made by Irène Curie-Joliot and F. Joliot[132, 133], that when a transmutation is provoked by particle bombardment the nucleus so formed may have a metastability comparable with that of the naturally radioactive atoms. Before this discovery the emission of a particle by a bombarded nucleus was always assumed to take place within an immeasurably short interval after the impact of the disrupting particle, although probably the experimentally established upper limit of the interval between the two events was not less than 1/100th sec., as evidenced, for instance, by Blackett's photographs of the tracks of expelled protons. At any rate, in the case of artificial transmutation, no emission of particles had ever been detected after the cessation of the bombardment. Curie-Joliot and Joliot, having established that certain light elements emitted positrons when bombarded by α-particles from polonium, proceeded to investigate the minimum energy of the α-particles required to produce this effect in aluminium. In the course of the investigation an aluminium sheet, 1 mm. thick, was exposed for 10 min. to α-rays, and then placed before the window, of 0·07 mm. aluminium foil, of a Geiger-Müller counter. A radiation was observed which decreased exponentially, with a half-period of 195 sec. Experiments made with the Wilson chamber established that the particles emitted were positrons, and the following equations were taken to represent the process for the bombardment:

$$^{27}_{13}\mathrm{Al} + {}^{4}_{2}\mathrm{He} = {}^{30}_{15}\mathrm{P} + {}^{1}_{0}n,$$

forming an unstable phosphorus isotope, which then breaks down according to

$$^{30}_{15}\mathrm{P} = {}^{30}_{14}\mathrm{Si} + {}^{+}_{\epsilon}.$$

They proceeded to establish the phenomenon for boron and magnesium as well. The period is independent of the energy of the α-rays used for bombardment. In the case of aluminium and boron only positive electrons are emitted, but in the case of magnesium both positive and negative electrons are detected, probably due to two different isotopes, which can be formed thus:

$$\left.\begin{aligned} {}^{24}_{12}\mathrm{Mg} + {}^{4}_{2}\mathrm{He} &= {}^{27}_{14}\mathrm{Si} + {}^{1}_{0}n \\ {}^{27}_{14}\mathrm{Si} &= {}^{27}_{13}\mathrm{Al} + {}^{+}_{\epsilon} \end{aligned}\right\},$$

$$\left.\begin{aligned} {}^{25}_{12}\mathrm{Mg} + {}^{4}_{2}\mathrm{He} &= {}^{28}_{13}\mathrm{Al} + {}^{1}_{1}\mathrm{H} \\ {}^{28}_{13}\mathrm{Al} &= {}^{28}_{14}\mathrm{Si} + {}^{-}_{\epsilon} \end{aligned}\right\}.$$

The emitted electrons give in all cases a continuous spectrum, just as do natural β-rays, and the hypothesis of a neutrino is equally necessary in both cases. The actual formation of the isotopes postulated in the reactions typified by the above equations has been confirmed chemically. For instance, with boron the reactions are assumed to be

$$\left.\begin{aligned} {}^{10}_{5}\mathrm{B} + {}^{4}_{2}\mathrm{He} &= {}^{13}_{7}\mathrm{N} + {}^{1}_{0}n \\ {}^{13}_{7}\mathrm{N} &= {}^{12}_{6}\mathrm{C} + {}^{+}_{\epsilon} \end{aligned}\right\},$$

which means that the active element is an isotope of nitrogen. Accordingly, boron nitride, BN, was irradiated and immediately decomposed by sodium hydrate into ammonia and boric acid. The residue was found to be without activity, while the ammonia, which would contain any isotope of nitrogen, showed the same activity as activated boron does. Boron was particularly favourable for these experiments, as the half-period of the artificial radioactivity is comparatively long, namely 14 min., but the chemical effects have been established in the other cases as well.

The work has been rapidly followed up. Cockcroft, Gilbert and Walton[134] have produced radionitrogen, $^{13}_{7}N$, which Curie and Joliot obtained from the boron isotope by α-particle bombardment, by exposing a carbon target to a heavy stream of protons of 600 kV. energy. The half-life period which they found was 10·5 min. as compared to the 14 min. reported by Curie and Joliot. The difference may be real, and due to the formation of the radionitrogen nucleus in a different excited state in the two methods. They confirmed that the main part of the radiation, at any rate, consisted of positive electrons. We have already, then, two distinct methods of forming radionitrogen. Another interesting piece of work has been carried out by C. D. Ellis and W. J. Henderson[135], who have measured the artificial radioactivity of aluminium (or rather of radiophosphorus, as the active isotope $^{30}_{15}P$ is called) as a function of the energy of the activating α-rays, and established a very rapid increase of the yield of radiophosphorus with the energy. The details of this dependence are obviously of interest in connexion with the study of the potential barrier, which governs the penetration of the α-particle into the nucleus. They have also investigated the energy spectrum of the emitted protons.

Another type of artificial radioactivity, of great importance, has been announced by Fermi[136, 137], who had the brilliant idea of using neutrons to effect the transformation which, on account of their lack of charge, penetrate the nucleus with ease. The source of neutrons was the original Curie-Joliot-Chadwick one, viz. beryllium bombarded by α-particles, and took the form of beryllium powder enclosed in a glass tube with a large amount of radium emanation. The elements to be transformed were in the form of cylinders, which were first put round the source of neutrons, and afterwards round a Geiger counter. In this way a whole series of new radioactive elements were produced, with parent elements of a very wide range of atomic numbers, from fluorine to barium. In most cases, the half-life period was some minutes, but in the case of the radio element from arsenic it was 2 days, while for that from chlorine it was even longer. It is too early to attempt to discuss the many interesting effects that are in the course of publication, but that important results are to be expected in this field is clear. It may be noted that, with one exception, the particle emitted by the new radio elements is a positive or negative electron. If it had been an α-particle or proton, artificial radioactivity would have been detected long ago.

§ 12. THE COMPONENT PARTS OF THE NUCLEUS

We have spoken of protons and neutrons as components of nuclear structure, and have considered many well-established transmutations which involve the entry of these particles into nuclear structure, and their emission. In addition, we have considered, in these transmutations, the diplon and the α-particle. There is good evidence, as mentioned in § 6, that the diplon can be broken up into a proton and neutron, and it has not yet been found as an emitted particle in transmutation reactions, so that there is clearly no need to consider it as a fundamental particle. As for the α-particle, it has never been disrupted in collision, but the lithium and boron transmutations by protons and diplons, described in § 6, clearly shows the manufacture of at least one α-particle, since, for instance, of the three α-particles emitted by ^{11}B when bombarded by protons only two can possibly have existed ready-made in the boron nucleus. It is generally believed that the α-particle is made up of two protons and two neutrons. On account of its great stability it is convenient to speak of it as a separate entity, both inside and outside the nucleus, but there is little doubt of its composite character.

So far we have said nothing of the electron as a feature of nuclear structure, although in discussing β-ray spectra we have spoken of its emission from the nucleus, which would seem to imply that it formed a component part of the nucleus. Such, however, is not the present belief: it is held, rather, that the electron is manufactured at the moment of its emission. Fermi has, for instance, suggested that the electron is ejected from the nucleus by a quantum switch somewhat analogous to that which ejects the photon of X-ray and optical spectra from the extranuclear parts of the atom. When a neutron switches over to become a proton an ordinary negative electron is liberated: when a proton switches over to a neutron a positron is emitted. On this view we can, then, think of neutron and proton as two different internal quantum states of one and the same fundamental particle.

Since the electron was so long described as a component part of the nuclear structure cogent reasons seem to be required for the revision just outlined, and these are to hand. In this connexion, the nitrogen nucleus is of great significance. If this nucleus consists, as was long supposed, of 14 protons and 7 electrons, it will contain an odd number of particles, and according to the quantum rules for the composition of moments of momentum it should have a spin of an odd number of times $\frac{1}{2}.\frac{h}{2\pi}$, since the spin of each individual particle is $\frac{1}{2}.\frac{h}{2\pi}$. The measured spin of the nitrogen nucleus is, however, $\frac{h}{2\pi}=2.\frac{1}{2}\frac{h}{2\pi}$. This has been called the nitrogen paradox: a similar state of things arises with Li, for which the nuclear spin is 0.

It is clear that this difficulty is at once eliminated if we assume that the nitrogen nucleus consists of 7 protons and 7 neutrons, each being treated as an elementary particle of spin $\frac{1}{2}.\frac{h}{2\pi}$, for now we have an even number of particles. Another difficulty which vanishes if the nucleus be supposed to consist of protons and

neutrons only, with its attendant change in number of particles as compared with the proton-electron hypothesis, concerns the type of statistics which holds for the nucleus. It has been pointed out in § 9 that the type of statistics which the elementary particles in a nucleus obey can be deduced from the alternating intensity of band-spectra lines, and the same thing can be found by an analysis of the Raman effect of the rotation terms in certain band spectra[138]. According to this evidence the nitrogen nucleus should obey the Bose type of statistics. However, a general rule in quantum statistics[139] states that an aggregate of elementary particles obeys the Bose statistics if it is even in number, and the Fermi statistics if it is odd. Here again, then, a constitution of protons and neutrons reconciles various results, while protons and electrons lead to contradictions[140].

Finally, another argument against including the electron as a nuclear constituent is furnished by Dirac's electrodynamics, which leads to grave difficulties if we try to confine an electron in the limited nuclear space. A rough analogy, which must not be pressed, is offered by the classical treatment of the light quantum, which must travel, and cannot be confined to one spot.

There is little doubt, then, that in the light of the present theory it is more satisfactory to exclude the electron as a nuclear inhabitant, and to deal with protons, neutrons and that particularly stable structure of protons and neutrons, the α-particle.

REFERENCES

(1) Gamow. *Z. Phys.* **51**, 204 (1928).
(2) —— *Z. Phys.* **52**, 510 (1928).
(3) Gurney and Condom. *Phys. Rev.* **33**, 122 and 127 (1929).
(4) Geiger and Müller. *Phys. Z.* **29**, 839 (1928).
(5) —— —— *Phys. Z.* **30**, 489 (1930).
(6) —— —— *Z. Phys.* **48**, 624 (1928).
(7) Werner Schulze. *Naturwissenschaften*, **20**, 315 (1932).
(8) —— *Z. Phys.* **78**, 92 (1932).
(9) Hoffman. *Ann. Phys.*, Lpz., **142**, 1196 (1913).
(10) —— *Ann. Phys.*, Lpz., **80**, 779 (1926).
(11) Greinacher. *Z. Phys.* **36**, 364 (1926).
(12) Wynn-Williams. *Proc. roy. Soc.* **132**, 295 (1931).
(13) —— *Proc. roy. Soc.* **136**, 312 (1932).
(14) Bothe and Kohlhörster. *Z. Phys.* **56**, 751 (1929).
(15) Rossi. *Z. Phys.* **82**, 151 (1933).
(16) Johnson and Stevenson. *J. Franklin Inst.* **216**, 329 (1933).
(17) Geiger and Klemperer. *Z. Phys.* **49**, 752 (1928).
(18) Wilson. *Proc. roy. Soc.* **142**, 88 (1933).
(19) Rutherford, Ward and Wynn-Williams. *Proc. roy. Soc.* **129**, 211 (1930).
(20) Ellis. *Proc. roy. Soc.* **135**, 628 (1932).
(21) Cockcroft and Walton. *Proc. roy. Soc.* **136**, 619 (1932).
(22) Lawrence and Livingstone. *Phys. Rev.* **40**, 19 (1932).
23) For a more detailed account of Aston's work, with references, see his book, already quoted, where Dempster's work is also described. See also E. N. da C. Andrade, *Rayons Positifs et leur application à la question des isotopes*, Congrès International d'Electricité, **2**, 891, where a full list of references up to the end of 1931 is given.

(24) Eisenhut and Conrad. *Z. Elektrochemie*, **36**, 654 (1930).
(25) Lukanow and Schütze. *Z. Phys.* **82**, 610 (1933).
(26) Aston. *Proc. roy. Soc.* **144**, 8 (1934).
(27) Urey, Brickwedde and Murphy. *Phys. Rev.* **40**, 1 (1932).
(28) Washburn and Urey. *Proc. nat. Acad. Sci.*, Wash., **18**, 496 (1932).
(29) Lewis and Macdonald. *Amer. chem. J.* **55**, 1297 (1933).
(30) —— —— *J. chem. Phys.*, **1**, 288 (1933).
(31) Rutherford. *Nature*, Lond., **133**, 481 (1934).
(32) —— *Ann. Rep. chem. Soc.* **31** (1932).
(33) Darmois. *Le Deutérium*, Paris (1934).
(34) Skobelzyn. *Z. Phys.* **43**, 354 (1927).
(35) Anderson. *Science*, **76**, 238 (1932).
(36) —— *Science*, **77**, 432 (1933).
(37) Anderson and Millikan. *Phys. Rev.* **40**, 325 (1932).
(38) —— —— *Phys. Rev.* **41**, 405 (1932).
(39) —— —— *Phys. Rev.* **43**, 491 (1933).
(40) Blackett and Occhialini. *Proc. roy. Soc.* **139**, 699 (1933).
(41) Skobelzyn. *C.R. Acad. Sci.*, Paris, **195**, 315 (1932).
(42) Blackett and Occhialini. *Proc. roy. Soc.* **139**, 699 (1933) (p. 715).
(43) Anderson. *Science*, **77**, 432 (1933).
(44) Joliot and Irène Curie-Joliot. *C.R. Acad. Sci.*, Paris, **196**, 1581 (1933).
(45) Meitner and Philipp. *Naturwissenschaften*, **24**, 468 (1933).
(46) Chadwick, Blackett and Occhialini. *Proc. roy. Soc.* **144**, 235 (1934).
(47) Rutherford. *Proc. roy. Soc.* **97**, 374 (1920).
(48) Bothe and Becker. *Z. Phys.* **66**, 289 (1930).
(49) —— —— *Naturwissenschaften*, **18**, 705 (1930).
(50) —— —— *Naturwissenschaften*, **19**, 753 (1931).
(51) Joliot and Irène Curie-Joliot. *C.R. Acad. Sci.*, Paris, **194**, 273, 708 (1932).
(52) Chadwick. *Nature*, Lond., **129**, 312 (1932).
(53) —— *Proc. roy. Soc.* **136**, 692 (1932).
(54) Feather. *Proc. roy. Soc.* **136**, 709 (1932).
(55) de Broglie, Leprince-Ringuet, Thibaud, *et al. C.R. Acad. Sci.*, Paris, **194**, 1647 (1932).
(56) —— —— —— —— *C.R. Acad. Sci.*, Paris, **195**, 88 (1932).
(57) —— —— —— —— *C.R. Acad. Sci.*, Paris, **195**, 665 (1932).
(58) Fermi. *Congrès internat. d'Elec.* **2**, 981 (1932).
(59) Meitner and Philipp. *Naturwissenschaften*, **20**, 929 (1932).
(60) Rasetti. *Naturwissenschaften*, **20**, 252 (1932).
(61) Auger. *C.R. Acad. Sci.*, Paris, **194**, 877 (1932).
(62) Chadwick. *Proc. roy. Soc.* **142**, 1 (1933).
(63) Ladenburg. *Phys. Rev.* **45**, 224 and 495 (1934).
(64) Langer. *Phys. Rev.* **45**, 137 (1934).
(65) Cockcroft and Walton. *Proc. roy. Soc.* **137**, 229 (1932).
(66) —— —— *Proc. roy. Soc.* **144**, 704 (1934).
(67) Dee and Walton. *Proc. roy. Soc.* **141**, 733 (1933).
(68) Lawrence, Livingstone and White. *Phys. Rev.* **42**, 150 (1932).
(69) Henderson. *Phys. Rev.* **43**, 98 (1933).
(70) Cockcroft and Walton. *Proc. roy. Soc.* **144**, 704 (1934).
(71) Lawrence, Lewis and Livingstone. *Phys. Rev.* **44**, 55 (1933). See also various short notices in the *Phys. Rev.* for 1933 and 1934.
(72) Oliphant and Rutherford. *Proc. roy. Soc.* **141**, 259 (1933).
(73) Oliphant and Kinsey. *Proc. roy. Soc.* **141**, 722 (1933).
(74) Lewis, Livingstone and Lawrence. *Phys. Rev.* **44**, 55 (1933).
(75) Oliphant, Kinsey and Rutherford. *Proc. roy. Soc.* **141**, 722 (1933).
(76) Oliphant and Rutherford. *Proc. roy. Soc.* **141**, 259 (1933).
(77) Kirchner. *S. B. bayer. Akad. Wiss.* p. 129 (1933).

(78) Dee and Walton. *Proc. roy. Soc.* **141**, 733 (1933).
(79) Oliphant, Harteck and Rutherford. *Proc. roy. Soc.* **144**, 692 (1934).
(80) Latimer and Young. *Phys. Rev.* **44**, 690 (1933).
(81) Rutherford *et al. Nature*, Lond., **133**, 413 (1934).
(82) Dee. *Nature*, Lond., **133**, 564 (1934).
(83) Feather. *Proc. roy. Soc.* **136**, 709 (1932).
(84) Harkins *et al. Phys. Rev.* **44**, 529 (1932), and subsequent notes.
(85) Rutherford, Ward, Wynn-Williams, Lewis and Bowden. *Proc. roy. Soc.* **129**, 211 (1930).
(86) —— —— —— —— —— *Proc. roy. Soc.* **131**, 684 (1931).
(87) —— —— —— —— —— *Proc. roy. Soc.* **133**, 351 (1931).
(88) —— —— —— —— —— *Proc. roy. Soc.* **136**, 349 (1932).
(89) —— —— —— —— —— *Proc. roy. Soc.* **139**, 617 (1933).
(90) Nimmo and Feather. *Proc. roy. Soc.* **122**, 668 (1929).
(91) Philipp and Donat. *Z. Phys.* **52**, 759 (1929).
(92) Rosenblum. *C.R. Acad. Sci.*, Paris, **188**, 1401 and 1549 (1932).
(93) —— *C.R. Acad. Sci.*, Paris, **194**, 967 (1931) and 1154 (1932).
(94) —— *J. Phys. Radium*, **1**, 438 (1930).
(95) —— *J. Phys. Radium*, **2**, 309 (1931).
(96) Rutherford *et al. Proc. roy. Soc.* **139**, 617 (1933).
(97) Ellis. *Proc. roy. Soc.* **138**, 318 (1932).
(98) Ellis and Mott. *Proc. roy. Soc.* **139**, 369 (1933). For earlier work see the book of Rutherford, Chadwick and Ellis.
(99) Thibaud. Thèse, Paris (1925).
(100) Ellis. *Proc. roy. Soc.* **129**, 180 (1930).
(101) Rutherford and Ellis. *Proc. roy. Soc.* **132**, 667 (1931).
(102) Gamow and Houterman. *Z. Phys.* **52**, 496 (1928).
(103) Ellis and Mott. *Proc. roy. Soc.* **139**, 369 (1933).
(104) Gamow. *Proc. roy. Soc.* **146**, 217 (1934).
(105) Bothe and Fränz. *Z. Phys.* **49**, 1 (1928).
(106) Pose. *Phys. Z.* **30**, 780 (1929).
(107) —— *Phys. Z.* **31**, 943 (1930).
(108) —— *Phys. Z.* **35**, 633 (1934).
(109) —— *Z. Phys.* **67**, 194 (1931).
(110) Chadwick and Constable. *Proc. roy. Soc.* **135**, 48 (1932).
(111) Chadwick, Constable and Pollard. *Proc. roy. Soc.* **130**, 463 (1931).
(112) Steudel. *Z. Phys.* **77**, 139 (1932).
(113) de Broglie and Leprince-Ringuet. *C.R. Acad. Sci.*, Paris, **193**, 132 (1931).
(114) Haxel. *Z. Phys.* **83**, 323 (1933).
(115) —— *Z. Phys.* **88**, 346 (1934).
(116) Geiger and Klemperer. *Z. Phys.* **49**, 752 (1928).
(117) Gurney. *Nature*, Lond., **123**, 565 (1929).
(118) Pauli. *Naturwissenschaften*, **12**, 741 (1924).
(119) Schüler *et al. Naturwissenschaften*, **21**, 674 (1933).
(120) Fermi and Segrè. *Z. Phys.* **82**, 729 (1933).
(121) Goldsmit. *Phys. Rev.* **43**, 636 (1933).
(122) Schüler and Brück. *Z. Phys.* **56**, 291 (1929).
(123) Schüler and Westmeyer. *Z. Phys.* **82**, 685 (1933).
(124) Schüler and Jones. *Naturwissenschaften*, **20**, 171 (1932).
(125) de L. Kronig. *Band Spectra*, Cambridge (1930).
(126) de L. Kronig and Frisch. *Phys. Z.* **32**, 457 (1931).
(127) Heitler and Herzberg. *Naturwissenschaften*, **17**, 673 (1929).
(128) Ornstein and van Wijk. *Z. Phys.* **49**, 315 (1928).
(129) —— —— *Z. Phys.* **59**, 313 (1930).
(130) Sergant. *Proc. Camb. phil. Soc.* **28**, 538 (1932).
(131) Meitner and Orthmann. *Z. Phys.* **69**, 143 (1930).

(132) Irène Curie-Joliot and Joliot. *C.R. Acad. Sci.*, Paris, **198**, 254 (1934).
(133) —— —— *J. Phys. théol. appl.* **5**, 153 (1934).
(134) Cockcroft, Gilbert and Walton. *Nature*, Lond., **133**, 328 (1934).
(135) Ellis and Henderson. *Proc. roy. Soc.* **146**, 206 (1934).
(136) Fermi. *Nature*, Lond., **133**, 757 (1934).
(137) —— *La Ricerca Scientifica*, **5**, 282 (1934).
(138) Rasetti. *Z. Phys.* **61**, 598 (1930).
(139) Ehrenfest and Oppenheimer. *Phys. Rev.* **37**, 333 (1931).
(140) The ideas here expressed seem to have arisen in more than one quarter after the discovery of the neutron. In particular, Heisenberg, *Z. Phys.* **77**, 1 (1932), **80**, 587 (1933), may be consulted.

ELECTRIC AND MAGNETIC MEASUREMENTS

By L. HARTSHORN, D.Sc.

§ 1. GALVANOMETERS

The limit of sensitivity. Galvanometers of so great a sensitivity have been made in recent years that the theoretical limit of sensitivity may be said to have been reached. The sensitivity attainable with any galvanometer depends entirely on the steadiness of its zero. In many galvanometers the effects of external mechanical vibration set the limit, but by the use of a stable mounting and the accurate balancing of the moving system, the effects of external vibration may be made so small that they are less than those of thermal agitation in the galvanometer system itself. This agitation is of course quite beyond our control (except by lowering the temperature), and therefore sets a natural limit to the precision obtainable with any instrument depending on a moving system (including galvanometers of all types). The agitation may be considered to arise from the bombardment of the moving system by the molecules of the atmosphere, or the random movements of the free electrons in the coil system of the galvanometer, which constitute small currents and therefore deflect the instrument; but whatever the cause, if the system is in equilibrium at an absolute temperature T there will be associated with each of its degrees of freedom kinetic energy of mean value $kT/2$ where k is Boltzmann's constant. A deflection θ corresponds to potential energy $\frac{1}{2}a\theta^2$, where a is the restoring couple per unit angle. Thus if $\bar{\theta}$ is the root-mean-square deflection due to thermal (Brownian) agitation, we have $a\bar{\theta}^2 = kT$ from which $\bar{\theta}$ may be estimated. Although a is not usually known, it may be evaluated in terms of known constants of the galvanometer, by making use of the ordinary equation of motion of the instrument. In the particular case of a moving-coil instrument of period t, critically damped by induction in the galvanometer circuit, which is of total resistance R, Ising[1] has shown that the value of $\bar{\theta}$ is equivalent to a current of $\sqrt{(\pi kT/Rt)}$ or a voltage of $\sqrt{(\pi kTR/t)}$, and if for example R is 50 Ω. and $t = 2$ sec. these become 10^{-11} A. and 5×10^{-10} V. respectively. Photographic records of galvanometer deflections obtained by Moll and Burger by the use of a moving-coil galvanometer and thermo-relay showed fluctuations of this order, and subsequently Orstein, Burger and collaborators[2] showed that the fluctuations were diminished by about the calculable amount when a coil included in the galvanometer circuit was cooled by liquid air. Thermal agitation of electrons in this coil was thereby diminished, and since it formed part of the galvanometer system, there was a corresponding diminution of the Brownian movement of this system. In order to calculate the diminution it was of course necessary to extend the theory to a system at two temperatures T_1 and T_2 instead of the single temperature T.

Moving-coil galvanometers. The Moll galvanometer has been known for some years as one of the most sensitive of moving-coil instruments. Moll and Burger[3]

have recently increased its sensitivity by adding to the moving system a needle of paramagnetic material, which is placed in a homogeneous magnetic field, and set in such a direction that the couple exerted on it by this field opposes the restoring force due to the suspension. In one type of instrument the needle is placed inside the moving coil and is directly influenced by the galvanometer magnet. In another the needle is set in an auxiliary field of variable strength, which thus provides a convenient means of adjusting the sensitivity. A voltage sensitivity of 1 mm. at 1 m. for 9×10^{-8} V., and a current sensitivity of 1 mm. at 1 m. for 3×10^{-10} A. are obtained, both with a period of 5 sec.

The limit of sensitivity set by the Brownian movements of the suspended system was first reached by the use of two Moll galvanometers and a thermal relay in cascade. G. Ising[4] in 1932 reached this limiting sensitivity by means of a single galvanometer with optical magnification. It is a moving-coil instrument, embodying several of the features of the Moll and Zernike instruments, but the moving system carries a pointer instead of the usual mirror, and the movements of the tip of this pointer are observed with a magnification of 2000 by means of an ultramicroscopic arrangement. For this purpose the whole galvanometer has been made so small that it can be mounted on the table of an ordinary microscope. The moving coil is of circular form, and is suspended between two short quartz fibres, which are kept under considerable tension by means of springs. By means of this construction it is possible to obtain a very high degree of symmetry, which is essential in order to free the system from disturbances due to external vibration. The current passes to the coil through two very fine silver strips. The coil swings in the radial magnetic field between a spherical iron core and the pole pieces of the magnet. The pointer attached to the moving system consists of a quartz fibre 32 mm. long, its outer end being drawn down so that the last 3 mm. is of diameter 0·017 mm. In order to obtain sufficiently good definition for accurate measurements with a magnification as high as 2000, the tip of the pointer was not observed directly, but sharp defraction fringes were derived from light passing through it, by making use of the Fraunhofer diffraction phenomenon. Light from an illuminated slit was brought to a focus near the pointer, and passed through it to an objective with a central stop, which cut off the direct beam, but transmitted the diffracted light through the two openings on either side of the stop. The fringes thus produced were observed by a high power microscope, and even after a 2000-fold magnification they were only 1 mm. wide, and therefore formed satisfactory reading marks. Any movement of the pointer produced a corresponding movement of the fringes. The agitation of the fringes when the galvanometer was "at rest" was recorded photographically, and it was found that the measured mean deviation of the moving system was only 5 per cent greater than the value calculated on the basis of the theorem of equipartition of energy. Thus practically all sources of disturbance other than those of the thermal agitation must have been eliminated.

In order to reach the limiting sensitivity directly with a galvanometer, the restoring force must be very small (0·14 dyne-cm. in the instrument described) and practical considerations, such as possible sizes of wire, limit the resistances of such

galvanometers to a certain range, say 3 to 300 Ω. Ising has shown however that it is possible to reach the limiting sensitivity for a galvanometer of a different order of resistance with a larger controlling force by the use of a transformer galvanometer in conjunction with one of the sensitive instruments described above. The transformer galvanometer is a galvanometer with two moving coils rigidly coupled together, one of the resistance required (e.g. a very high resistance for the measurement of ionization currents) and the other of about the same resistance as the galvanometer of limiting sensitivity, to which it is connected. It is then shown that a deflection of the transformer galvanometer, which is a certain multiple of its mean Brownian deviation, causes a deflection of the sensitive galvanometer, which is the same multiple of its own Brownian deviation, and since the latter is directly observable, the limiting sensitivity of the first galvanometer is also reached. The deflection of the second galvanometer of course only occurs when the current through the first galvanometer changes, e.g. when this current is started or stopped. The same principle of deflection magnification may be applied to instruments other than galvanometers.

Moving-magnet galvanometers. Although astatic moving magnetic galvanometers were for many years the most sensitive available, they are generally regarded as instruments to be avoided whenever possible, since the astaticism is never perfect, and disturbances of the zero due to stray magnetic fields are often so serious as to make the instrument unworkable. The use of shields of material of very high permeability (e.g. mumetal) is effective up to a point, since the magnetic field inside the instrument is greatly reduced by their use, but since the shield cannot possibly be complete, it may produce small irregularities of field, for which the magnet system is not compensated. The moving-magnet galvanometer, even when shielded, is therefore in general greatly inferior to the moving-coil instrument, and since the latter has reached the limit of sensitivity, it might be thought that the former is obsolescent. There remains however the questions of difficulty of construction and cost, and a new moving-magnet galvanometer of high sensitivity due to L. Meylan[5] is not without interest. The astaticism has been improved in two ways. Four systems of magnets (two astatic pairs, symmetrically arranged along the suspension) are used instead of the usual two, so that the system does not demand a uniform field for astaticism, but only one of uniform gradient. In addition the final adjustment of astaticism can be made electrically, while the moving system is freely suspended in its final position. The magnets consist of very thin discs of cobalt-chromium steel (36 per cent cobalt), the axis of suspension passing through the centres and lying in the plane of the discs. Four pairs of coils are provided, one pair acting on each of the magnets. For the purpose of compensation an additional disc of cobalt-chromium steel is mounted on the suspension, with its plane at right angles to the axis. This disc may be magnetized in two directions at right angles, by means of two pairs of auxiliary coils. The necessary current is passed through these coils for very short periods only (of the order of 0·01 sec. controlled by a Helmholtz pendulum arrangement) so that clamping of the moving system is unnecessary. By successive trials the direction and intensity of magnetization of the disc may be made to balance

the residual moment of the system of four magnets. A quartz fibre suspension is employed, and the damping is mainly electromagnetic, due to induction in the coils and small pieces of electrolytic copper which may be placed in suitable positions near the magnets. The mean variations of zero with a period of 8 sec. were 0·1 mm. at night and 0·25 mm. in the day time, and the current sensitivity for this condition was 1 mm. for 4×10^{-11} A. with a coil resistance of about 2000 Ω. With a coil system of resistance 0·5 Ω. a voltage sensitivity of 6×10^{-10} V. for 1 mm. at 1 metre was obtained with a period of 10 sec.

Thermal amplifiers or relays. The thermoelectric relay, first devised by W. Hamilton Wilson and subsequently developed with great success by Moll and Burger, is now widely used when the greatest possible sensitivity is required. It consists essentially of two similar thermocouples connected in opposition to a galvanometer (the secondary galvanometer). A powerful beam of light reflected from the mirror of the primary galvanometer is concentrated at a point midway between the two junctions, so that both are equally heated and no deflection of the secondary galvanometer occurs. Any deflection of the primary galvanometer displaces the spot of light towards one junction, thereby causing unequal heating of the two junctions, and a deflection of the secondary galvanometer. The magnitude of the secondary deflection is easily controlled by altering the intensity of the beam of light; 100-fold amplification may be obtained, and the limit of sensitivity may be reached by means of two moving-coil galvanometers, which need not be of the most sensitive type. The time of operation is 2·5 sec.

A thermal relay of an entirely different type has been introduced by H. Sell[6]. It is described as a mechanically controlled bolometer, and depends for its action on the hot-wire principle. The bolometer consists of four very narrow coils of nickel wire (diameter of wire 0·02 or 0·03 mm.) arranged in the form of a Wheatstone bridge. The current in the bridge network is sufficient to raise the wires to a high temperature. Two of these coils in adjacent arms of the bridge are placed just behind two very narrow slits (0·1 mm. wide) so that streams of air issuing from the slits impinge on the wires and exert a cooling action. Two vanes carried by the moving system of a sensitive instrument are arranged so that their motion tends to open and close the slits more or less completely and so regulate the two streams of air. The positions of the slits are adjusted with reference to the vanes until normally the Wheatstone bridge is balanced, but a slight displacement of the vanes increases the cooling action on one wire and diminishes it on the other, and therefore disturbs the balance, and gives a comparatively large deflection of the instrument acting as detector in the bridge circuit. It will be realized that the power supplied to the bolometer network may be made comparatively large, and that therefore the output-power of the system is relatively very large. It is claimed that the power amplification may be as high as 10^{11}. Unsteadiness of the air streams appears likely to be the chief limitation. These are obtained by means of a small blower, one form of which consists essentially of a telephone receiver with a large diaphragm forming one wall of a very narrow air-cell. The telephone is operated by alternating current from the mains, and as the diaphragm vibrates, large changes of pressure are produced in the

air-cell owing to its very small volume. The high velocity of the diaphragm and the large changes of pressure produced make it possible to obtain suitable streams of air from outlet tubes communicating with the air cell. The whole apparatus, which is made commercially, is quite compact. The time of operation is given as less than 1 sec. and displacements of the vanes as small as $0{\cdot}5 \times 10^{-4}$ mm. are measurable. The amplifier may of course be applied to any instrument with moving parts, and is also suitable for all measurements involving the determination of small displacements, e.g. thermal expansion. Applied to galvanometers it has given a voltage sensitivity of 10^{-8} V., and a current sensitivity of 10^{-11} A., with a time of indication of 1 sec., the readings being made directly on pointer instruments.

§ 2. THERMIONIC AMPLIFIERS FOR THE MEASUREMENT OF CURRENT AND VOLTAGE

A very large amount of work has been done in recent years on the application of the thermionic valve to the measurement of very small currents and voltages, especially in connexion with work on photo-electric cells, ionization currents, and measurements of H ion concentration. The simplest possible circuit for such work is shown in figure 1. The voltage or current to be measured is applied to the grid

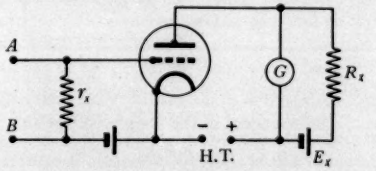


Figure 1.

circuit of the valve, i.e. to the input terminals A and B of figure 1. The result is a change in the value of the current in the anode circuit, and this is observed as a deflection of the galvanometer G. The reading of the galvanometer is adjusted to zero for zero input current or voltage, by shunting it with an auxiliary battery E_x and adjustable resistance R_x. The voltage sensitivity of this arrangement may be expressed as follows:

$$\text{Voltage sensitivity} = \frac{\delta I_G}{\delta e_g} = \frac{G_m}{1 + R_G\,(1/R_i + 1/R_x)} \qquad \text{.....(1),}$$

where δI_G is the change of galvanometer current caused by a change of input voltage δe_g; G_m is the mutual conductance of the valve, R_i its internal resistance, R_G the resistance of the galvanometer, and R_x the auxiliary resistance. The current sensitivity is given by

$$\text{Current sensitivity} = \frac{\delta I_G}{\delta i_{AB}} = \frac{G_m}{1 + R_G\,(1/R_i + 1/R_x)} \times \frac{r_x}{1 + r_x/r_i} \qquad \text{.....(2),}$$

where r_x is the external and r_i the internal resistance in the grid circuit of the valve and i_{AB} is the input current.

Full advantage of the magnifying properties of the valve can only be taken if for the galvanometer G we may use a very sensitive instrument, but it will be obvious that small unavoidable fluctuations of the voltages of the filament and h.t. batteries will always cause fluctuations of the anode current, and that the reading of G is likely to be very unstable if a sensitive instrument is used. Much ingenuity has been exercised in devising circuit arrangements in which the reading of the indicating instrument is independent of battery fluctuations. A few typical examples will be considered.

Bridge circuits. An alternative method of reducing the galvanometer current to zero for zero input voltage is shown in figure 2. This may be regarded as a

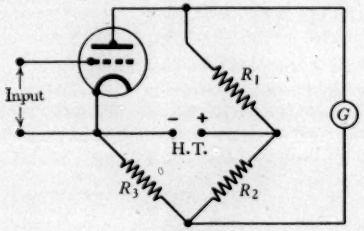


Figure 2.

Wheatstone's bridge circuit, and as in the case of the ordinary resistance bridge, the galvanometer current is independent of the supply voltage (h.t. in this case) when

$$R_i R_2 = R_1 R_3 \qquad \ldots\ldots(3).$$

By a suitable choice of grid and h.t. potentials this condition may be satisfied and also the galvanometer current may be reduced to zero, and the arrangement then becomes a voltmeter which is compensated for small fluctuations of the voltage of the h.t. battery, i.e. fluctuations which are not large enough to affect the value of R_i, say 1 to 3 per cent. This advantage is gained at the expense of a certain loss of sensitivity. If, as is usual, $R_1 = R_2$ and therefore $R_3 = R_i$, the sensitivity of this arrangement is given by

$$\frac{\delta I_G}{\delta e_g} = \frac{G_m}{2 + R_G\,(1/R_i + 1/R_1)} \qquad \ldots\ldots(4),$$

and Nottingham[7] has shown that if the same galvanometer and valve are used for the two circuits, and the circuit is arranged for critical damping of the galvanometer in each case, the first circuit has twice the sensitivity of the bridge circuit. However sensitivity is of no value unless it is accompanied by stability, and the advantages of the bridge circuit are considerable, especially when it is given the form of figure 3, in which the resistance R_3 is replaced by a second valve similar to the first. This second valve is merely a dummy whose function is to compensate for changes in the

operating valve. If the two valves were exactly similar in all their characteristics it is obvious that compensation would be perfect, but the matching of valves is a rough process only, and what may be regarded as second order fluctuations will occur. Closer approximation to true compensation may however be made if the grid and filament voltages of the two valves are supplied from a common source, but are arranged to be separately adjustable by means of potentiometers or series resistances. Thus Wynn-Williams[8, 9] showed that by having a small adjustable resistance in series with each filament and then connecting them in parallel across a common battery, it is possible to make the galvanometer reading independent of the voltage of this battery (within certain limits) by adjustment of the two series resistances. MacFarlane[10, 11] has carried the process a stage farther. He uses one battery for supplying both filaments and both grids. A potentiometer is connected across this battery and the various grid and filament terminals are connected to tapping points on the potentiometer. By suitable adjustments of these tapping points, the circuit is compensated for changes of both grid and filament potentials, in addition to the compensation for h.t. battery changes described above. The conditions necessary for compensation have been worked out, but in practice it is always done by trial and error methods, i.e. an arbitrary setting is made, and then a definite change in the appropriate battery voltage is made and the corresponding galvanometer deflection noted. An adjustment is then made, and a second reading obtained for the same change of voltage. If the reading is smaller and in the same direction, the adjustment has been made in the right direction, and so on. Full circuit details are given in the original papers.

The electrometer valve. In the preceding discussion it has been assumed that the valves described are ordinary triodes. The choice of valve will be determined by the kind of measurement to be made. For example Wynn-Williams[8, 9] pointed out that the maximum voltage sensitivity obtainable with his arrangement is proportional to the mutual conductance G_m or μ/R_i where μ is the voltage factor and R_i the internal anode circuit resistance of the valve. He therefore states that valves of large G_m, i.e. power valves, should be used. The circuit then forms a sensitive voltmeter of high but not negligible resistance. Its actual resistance (if no grid leak is used) is the input resistance or the internal grid-circuit resistance of the valve, i.e. $\delta e/\delta i$ or r_i where e and i are the grid potential and current respectively. (It is to be noted that we are only concerned with changes in e and i and not their actual values, whether the voltages to be measured are direct or alternating.) For many purposes an input resistance equal to that of a power valve would be too low. For example, the current sensitivity is limited by the value of r_i, as may be seen from equation (2). If therefore we require a device for the measurement of very small currents we must find a valve of high input resistance, i.e. of low grid current. If the grid current can be made negligibly small, the valve circuit becomes an electrometer, the grid exercising a purely electrostatic control of the anode current. Valves closely approximating to this ideal have recently been produced, and it is claimed that they have gone far towards making other electrometers obsolete.

The first description of an electrometer valve was given by Metcalf and

Thompson[12]. They point out that even when the grid of a valve is negatively charged, grid current is produced in many ways; by leakage over the glass container and insulating supports; by the attraction of positive ions, either emitted by the filament or resulting from the ionization of the residual gas by the electron current; by electron-emission from the grid due either to the heat and light from the filament or to soft X-rays from the anode produced by the normal anode current. Steps were taken to reduce to a minimum all these components of grid current. The grid connexion was brought out of the top of the bulb so as to make the leakage path as long as possible, and the valve was used in an enclosure dried by phosphorus pentoxide; the highest possible vacuum was used and all the electrode potentials were kept below 8 volts so as to avoid ionization. A space-charge grid, positively charged, was placed between filament and control grid so as to repel any positive ions leaving the filament. The filament was of thoriated tungsten with two legs in parallel to keep down the working voltage. With a control grid potential of -4 V. the grid current

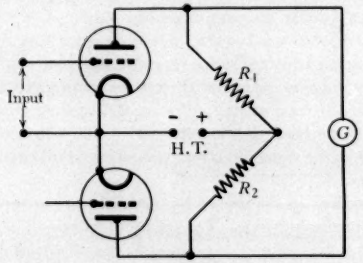


Figure 3.

was found to be of the order of 10^{-15} A., the input resistance $10^{16}\,\Omega$., and the input capacity $2{\cdot}5\,\mu\mu$F. Valves showing a similar performance are now produced by several of the leading firms.

Current measurement. Several observers[13, 14, 15] have described the measurement of very small currents by means of such valves, and it is claimed that they give a sensitivity exceeding that of any form of electrometer except the Hoffmann and with much greater ruggedness and stability. Thus the single valve circuit of figure 1 can be used to measure currents of 10^{-13} A. with an accuracy of 1 per cent ($r_x = 10^{10}\,\Omega$. and the galvanometer gives a deflection of 1 mm. for 5×10^{-10} A.). If r_x is omitted (the control grid is then "free") and the current measured by the rate of rise of grid potential, still greater sensitivity is obtained and currents of 10^{-16} A. have been measured in this way, but for currents of 10^{-14} A. or less the balanced circuit of figure 3 is preferable. When $r_x = 10^{10}\,\Omega$. this arrangement has been used with a sensitivity of 4×10^{-16} A./mm. Still greater sensitivity is obtained by using no grid leaks—i.e. removing r_x. Both grids are then free, and will change in potential at a rate depending on the grid current in the two valves and the grid capacities. A small

variable air condenser is connected between grid and filament of one valve and this is adjusted until the rate of rise of potential is the same for the two and the galvanometer therefore remains at zero. The current to be measured is then allowed to charge one grid, and the resulting rate of change of galvanometer deflection is measured. The only limit to the sensitivity is set by the fluctuations in the circuit. These are minimized by the use of good insulation (amber or sulphur at important points) with guard rings where practicable, shielding, batteries of large capacity, and good switch contacts (platinum or gold) and by switching on the valves an hour or two before the measurements. A current of 5×10^{-18} A. (30 electrons per sec.) has been measured in this way.

If a two-stage amplifier is used the sensitive galvanometer may be replaced by a more robust instrument, and many circuits of this form have been described for various purposes. Owing to the complex character of the grid current of a valve, the internal grid resistance of the valve may be either positive or negative, depending on the grid potential. Nottingham[7] pointed out that by making r_i negative, the current sensitivity of an amplifier may be considerably increased. This will be obvious from equation (2). MacDonald and others[16, 17] have used a two-stage amplifier, in which the value of r_i for the second valve is made negative. Since this is the value of R_x in equation (1) for the first valve, the voltage sensitivity of the arrangement is increased in a similar way. Two screen-grid valves were found to give a sensitivity of 30,000 μA./V. Another device for obtaining high voltage sensitivity without using a delicate detecting instrument is described by J. Roulleau[21]. The voltage to be measured is made to control the amplitude of an auxiliary alternating current, which may readily be amplified, then rectified, and measured with a milliameter. The alternating current (of audio frequency) is produced by coupling through a transformer, the anode circuit and one grid circuit of a two-grid valve. The potential of the other grid then determines the resistance of the valve and therefore the amplitude of the oscillations. This grid and the filament are therefore the input terminals of the voltmeter. With two stages of amplification following the tetrode the instrument would detect voltage changes of 2×10^{-5} V. The response was rather slow (0·25 sec.), so that the instrument could not be used for alternating current.

Compensation of two-grid valves. The electrometer valve described, like the ordinary screen-grid valve, has two grids. Such valves may be used in the bridge circuit already described but the presence of the second grid allows a simpler method of compensation for battery fluctuations to be used. If the potentials of the filament, positive grid and negative grid are all supplied by a common potentiometer and battery, it will be obvious that battery fluctuations will affect all the potentials simultaneously and by proportional amounts, and since the changes in the positive grid potential act on the emission current in the opposite direction to the others, the changes may be made self-compensating by a suitable choice of potentiometer tapping points. Circuits acting on this principle have recently been described by several authors[14, 18]. That of Soller shown in figure 4 may be regarded as typical. The single valve circuit shown will measure currents down to 5×10^{-14} A.

by the steady-deflection method. If the resistance r_x is omitted the arrangement becomes an electrometer which can be used to measure currents down to 10^{-15} A.

E.m.f. measurements on cells of high resistance. The electrometer valve with its very small grid current has been shown to be specially useful for the measurement of the e.m.f.'s of cells of very high resistance or those which are very easily polarized. Ellis and Kiehl[19] have shown that a valve amplifier using ordinary valves may also be used for such work if the input current is limited by placing a small air condenser of very high insulation resistance in series with the input terminals. Good results were obtained with cells with added resistance up to 500,000 MΩ. A potentiometer measurement is made with the amplifier and series condenser as detector of the balance point. Burroughs and Ferguson[20] point out that such methods of measure-

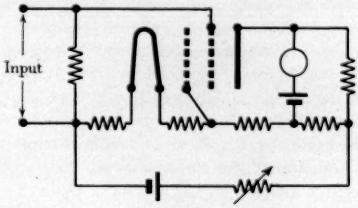


Figure 4.

ment are limited to cells of resistance less than about 10^{11} ohms, since even if the input capacity of the amplifier can be reduced to 1 $\mu\mu$F. (for an electrometer valve it is 3 $\mu\mu$F.) the time constant of the circuit with a cell of resistance 10^{12} ohms would be 1 sec. Thus on making an adjustment the observer must wait several seconds before its effect can be judged, and since perfect steadiness of the galvanometer is never obtained, the method becomes unworkable.

§ 3. ALTERNATING CURRENT GALVANOMETERS

The measurement of alternating currents with very great sensitivity is much more troublesome than the measurement of direct current, although as a result of the development of the thermionic amplifier very great sensitivity can be obtained for the detection of small alternating currents and voltages in null methods of measurement; for such currents and voltages may be magnified to almost any desired extent by the use of several amplifying valves in cascade, and the magnified current may then be detected by means of the vibration galvanometer or the telephone receiver. These instruments do not of course respond to the direct currents of the amplifier system, and there is therefore no need to adopt compensating devices, such as those discussed in the preceding section, in amplifiers for this purpose. The problem is thereby greatly simplified. But although very small alternating currents are easily detected, their measurement is a different matter. For this purpose the telephone and vibration galvanometer are obviously unsuitable.

For the direct measurement of alternating currents we have thermal instruments, dynamometers and moving-iron instruments, all instruments with a square-law scale. Of these only thermal instruments are suitable for high-frequency work, and indeed for the direct measurement of small alternating currents of all frequencies, the combination of heater, thermojunction and d.-c. moving-coil instrument has very largely superseded all other instruments. The heater and thermojunction are mounted in an evacuated glass bulb, the two being in thermal but not in electrical contact (the thermojunction is connected to the heater by means of a small bead of vitreous insulating material). This is of importance at high frequencies, for which the capacitance to earth of the d.-c. galvanometer, if directly coupled to the measuring circuit, might cause serious errors. Rather higher sensitivity can be obtained by having the junction in metallic contact with the heater but the difference is not great. Dynamometers and moving-iron instruments are not sufficiently sensitive to be of very great use for laboratory purposes, although the moving-iron instrument has been greatly improved by the use of the new nickel-iron alloys of very low coercive force.

The square-law instruments are of course all relatively insensitive for the measurement of very small currents, and in this respect they are much inferior to an ordinary d.-c. galvanometer. The advantages of this instrument would become available for the measurement of alternating current and voltage if a perfect rectifier could be found. Although existing rectifiers are by no means perfect, they have been so much improved in recent years that they are of considerable importance for measurement purposes.

Rectifiers of three types are now in use, thermionic, mechanical, and barrier-layer rectifiers consisting of a layer of semi-conducting material (cuprous oxide or selenium) in intimate contact with a metal plate. The last, the action of which is electronic, are also sometimes called "dry-plate" rectifiers to distinguish them from the older electrolytic rectifiers containing liquid, and partly chemical in their action, and crystal rectifiers the action of which was confined to certain sensitive spots. Neither the electrolytic nor the crystal rectifier was sufficiently stable to be used for measurement purposes.

The combination of thermionic rectifier and d.-c. galvanometer is the well-known thermionic voltmeter. Space does not permit any detailed consideration of the various circuits which are in use, but it is to be noted that zero stability is of the first importance, and this may be obtained by the use of the various compensated arrangements described in the previous section. These circuits may be used as sensitive thermionic voltmeters for a.-c. work; it is only necessary to adjust the voltages applied to the working valve, so that it is operating at a portion of its characteristic curve, where the curvature is large, so that rectification takes place. The thermionic voltmeter is specially valuable in that it may be used at any frequency. The ordinary forms have a square-law scale.

Mechanical rectifiers in the form of the rotating commutator and the vibrating reed have been used for many years. They are only applicable to comparatively low frequencies. The rotating commutator requires a synchronous motor and thus

becomes rather elaborate. The vibrating reed, which is much simpler, has recently been improved [22] so that it may be used in conjunction with a fairly sensitive d.-c. galvanometer. Thus an instrument with a full-scale reading of 10 μA. is a possibility, and it has the advantage of a linear scale, although, of course, it indicates mean value and requires a phase adjustment. H. Pfannenmüller [23] has described various circuits for obtaining this adjustment.

Copper-oxide rectifiers. The copper-oxide rectifier is the most widely used of the barrier-layer rectifiers. The rectifiers are made [24] by oxidizing discs of copper at a temperature of about 1040° C., so that the metal becomes coated with a layer of cuprous oxide. An outer layer of black oxide is also formed, but this is removed by dissolving in sodium cyanide solution. Rectification occurs at the boundary of the copper and cuprous oxide. The mother copper serves as one electrode, and the other generally takes the form of a lead disc pressed into contact with the outer surface of the oxide. In order to improve this contact, graphite may be rubbed into the surface of the oxide and a layer of tin foil may be placed between it and the lead, but no rectification occurs at this electrode.

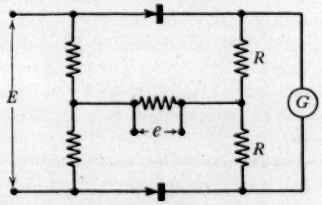


Figure 5.

In order to obtain full-wave rectification four such rectifiers are usually connected in the well-known bridge arrangement. The combination of such a rectifier with a d.-c. instrument makes a very convenient a.-c. instrument, and portable voltmeters, milliameters and microammeters working on this principle are now widely used. Instruments of this type will not however measure voltages less than about 0·025 V., since rectification no longer occurs when the applied voltage is less than this amount (at least the extent of the rectification becomes too small to be useful) and for the same reason very small currents cannot be measured with this device. This difficulty has been overcome by C. H. Walter [25] by polarizing the rectifier. Several arrangements are possible, the simplest being shown in figure 5. Two rectifiers are connected in a bridge network, to which is applied the small alternating voltage e to be measured, and also a comparatively large auxiliary voltage E of the same frequency as e. The circuit is so arranged that $(E+e)$ acts on one rectifier, and $(E-e)$ on the other. The corresponding rectified currents pass through two equal resistances R, across which a d.-c. galvanometer is connected. The current through this galvanometer is approximately proportional to the difference between the amplitudes of the vectors $(E+e)$ and $(E-e)$, since for voltages well above the threshold

value the characteristic curve of the rectifier is approximately linear. This difference of amplitudes is approximately equal to $2e \cos \alpha$, where α is the phase difference between the voltages E and e. Thus the deflection is almost independent of the magnitude of the polarizing voltage, which merely serves to increase the efficiency of the rectifier. It is to be noted that the instrument is sensitive to phase, and that the deflection changes sign when the direction of the voltage e is reversed. These are very useful properties in an instrument to be used as a detector in bridge measurements, and it is claimed that the use of a detector of this type makes the balancing of an a.-c. bridge as simple as that of a d.-c. one. The arrangement is applicable to rectifiers of other forms, and a similar device has been used by Pfannenmüller with the vibrating-reed rectifier. The copper-oxide type has however the advantage that it may be used at frequencies up to 10,000 c./sec.

The copper-oxide rectifier is effective at radio frequencies, but owing to the fact that the discs possess appreciable capacitance they must be regarded as possessing fairly large capacitance shunts, so that they become less sensitive as the frequency is increased. Special rectifiers of comparatively low capacitance have been produced for radio-frequency work, and they are much more satisfactory than the older crystal rectifiers, but they should be calibrated at the working frequency if this is higher than say 10,000 c./sec.

The differential thermionic voltmeter. A device with somewhat similar properties, but employing thermionic valves as the rectifiers, has been used by K. Kreielsheimer[26] for bridge measurements at radio frequencies. The circuit is that shown in figure 3, and it is used as a differential thermionic voltmeter. The two grids are connected to the detector terminals of the bridge, and the common filament terminal is connected to the middle point of the two arms which are to be balanced. Thus one valve gives a rectified anode-current approximately proportional to V_1^2, and the other gives one proportional to V_2^2, where V_1 and V_2 are the voltages across the two arms. The galvanometer reading is therefore proportional to $V_1^2 - V_2^2$, or $(V_1+V_2)(V_1-V_2)$, and when V_1 is nearly equal to V_2 the arrangement is very much more sensitive than a simple thermionic voltmeter would be, since this, if connected across the detector terminals, would give a reading proportional to $(V_1-V_2)^2$. The deflection, moreover, reverses its direction when the sign of V_1-V_2 changes.

Vibration galvanometers. In order to obtain the maximum sensitivity with a vibration galvanometer, the deflecting system must be tuned to resonance with the current supply. The resonance curve of a good galvanometer is very sharp so that a variation of 1 per cent in the frequency of the current may reduce the sensitivity by 100 per cent. In order to overcome this disadvantage without any great loss of sensitivity, Meissner and Adelsberger[27,28,29] have adopted the plan of connecting a condenser in series with the windings of the galvanometer, the value being chosen so that electrical resonance of the coil and condenser system also occurs in the working range. Thus two resonance curves (one mechanical and one electrical) are combined and the galvanometer sensitivity may be made nearly constant over a small range of frequencies about 10 c./sec. in extent. They have recently described an instrument specially designed for working on this principle. It is of the moving magnet type,

the magnet consisting of two very small flat plates of cobalt steel, one on either side of the suspension which is of platinum wire or strip held taut by a spiral spring. The plates, 5 mm. high and 1 mm. wide, are magnetized in the direction of their width. Since the material is magnetically very hard, this disposition gives the maximum ratio of magnetic moment to moment of inertia, and at the same time keeps the damping due to air resistance very small, thus giving high sensitivity and quick response. The moving magnet is located in the field produced by two electromagnets at right angles. One is an alternating current magnet, the windings of which carry the current to be measured, the alternating field produced deflecting the moving system. The other magnet is energized by direct current, and its field affects only the restoring force on the moving system, and thus controls its frequency of vibration. The tuning of the moving system is therefore carried out by adjusting the current supplied to the d.-c. magnet. The a.-c. magnet has a laminated core of large cross section, and conical pole tips, the air gap in which the moving magnet is suspended being only 2·5 mm. wide. The winding of this magnet has an inductance of one henry, so that it is not necessary to use inconveniently large condensers in order to satisfy the conditions for series resonance at ordinary working frequencies. Three types of instrument have been made to cover the frequency ranges 50 to 400 c./sec., 400 to 670 c./sec, and 670 to 800 c./sec. The energy relations of the system have been investigated theoretically, and applied to the consideration of suitable transformers and amplifiers for use with the galvanometer in order to obtain the maximum sensitivity in circuits of a given impedance.

§ 4. THE MEASUREMENT OF FREQUENCY

In few subjects has so much progress been made in recent years as in the measurement of the frequency of alternating currents or electrical oscillations. The subject became of considerable importance as soon as broadcasting became popular, since in order to minimize interference between the numerous stations it was essential that the frequency of each station should be controlled and measured within very close limits. Thus workers in many countries concentrated on the production of frequency standards of the highest possible accuracy, with the result that the accuracy was increased in a few years from a few parts in ten thousand to a few parts in ten million. An accuracy of this order is seldom required for the measurement of any frequencies other than those of high-frequency currents, but as the accuracy of radio-frequency standards increased, the possibility of using them as standards of time instead of the more usual pendulum clock was realized, and it now seems probable that clocks controlled by quartz oscillators vibrating at high frequency will prove to be more accurate than those of any other known form. One of the most important developments of the subject is the very wide application of the harmonic principle, as a result of which almost any desired multiple or sub-multiple of the frequency of a given standard may be obtained with an accuracy equal to that of the standard. Thus measurements of very high frequencies are easily made in terms of a low frequency standard, and *vice versa*. We therefore find tuning forks of

comparatively low frequency and quartz oscillators of high frequency used as alternative standards for the measurement of radio and audio frequencies.

The tuning-fork standard. A great improvement in the accuracy of the tuning fork as a standard of frequency was produced when Eccles showed how to maintain a tuning fork in continuous vibration by means of a thermionic valve circuit. Dye investigated the properties of the Eccles system, and showed in 1923 that by a suitable choice of capacitance and inductance in the circuit, a stability of frequency of 2 parts in 100,000 could be obtained with a tuning fork vibrating at 1000 c./sec., without a very elaborate design. When greater accuracy was called for the various factors causing variations of the frequency were studied in detail, and as a result a standard of much greater accuracy was developed. The temperature coefficient of frequency of an ordinary steel fork is of the order of one part in 10,000 per 1° C. and a great improvement in this respect was made by making the fork of elinvar, a nickel steel having a very low temperature coefficient of elasticity. The temperature coefficient of frequency was reduced to a few parts in a million per 1° C. The apparatus was then placed in a thermostat, which kept the temperature of the fork constant to a few thousandths of a degree, so that uncertainties due to temperature were made less than 1 part in ten million. The method of supporting the fork also required investigation. The fork itself must be firmly clamped by its stem, so that its position with reference to the associated magnetic circuit is definitely fixed. Any variation in position alters the coupling between the fork and the valve circuit, and so changes the frequency. The clamping must also be such that the smallest possible amount of vibration is transmitted to the base or lost by friction in the clamp, since such a dissipation of energy increases the decrement of the fork, i.e. diminishes its sharpness of resonance, and thereby makes its characteristic frequency less definite. The actual clamping area was therefore made as small as possible, and it was located as near as possible to a node of the system. The prongs of the fork were moreover carefully balanced so that the minimum amount of energy was transmitted to the mounting. The actual change of frequency on clamping and unclamping the mounting to a massive base was used as a criterion of this condition, and the final balancing was done by lightly filing the prongs of the fork at the ends to reduce the length, or near the base to reduce the stiffness, until this change of frequency had a minimum value. In this way the logarithmic decrement was reduced to about 0·0004. The frequency of the fork was also found to be affected by changes of atmospheric pressure, the change being of the order of 2 or 3 parts in a million per cm. of mercury. The apparatus was therefore placed in a sealed enclosure, provided with a manometer. In its final form the fork was found to have a frequency stability of about 3 parts in ten million over periods of a week or more. Such variations as occurred were largely due to fluctuations in the supply voltages of the valve circuit, and when these were maintained constant by hand adjustment, the frequency appeared to be constant to a few parts in 10^9 over periods of an hour. Forks of this design are now made commercially, and are widely used as frequency standards in wireless stations and for other applications where high accuracy is required. In forks of lower accuracy, Yanouchevsky of the firm of Belin has been able to dispense

with the thermostat by employing a system of temperature compensation. Included in the magnetic circuit of the fork is a piece of steel having a comparatively large temperature coefficient of permeability. Thus the magnetic flux in the air-gap between prongs and pole pieces varies with temperature and thereby causes changes of frequency which may be made to compensate the resultant change due to other causes. The compensating effect may have either sign, depending on whether the compensating element is in series or parallel with the part of the magnetic circuit including the air-gaps.

The luminous-quartz resonator. It has long been known that pieces of quartz crystal placed between electrodes to which an alternating e.m.f. is applied are thrown into mechanical vibration in virtue of the inverse piezo-electric effect. Cady showed in 1922 that, if such a piece of quartz is connected to an oscillating circuit, and the frequency of the electrical oscillations is adjusted to be equal to the natural frequency of the piece of quartz, the latter is thrown into mechanical resonance, which may be detected by its reaction on the oscillatory circuit (the direct piezo-electric effect). The mechanical resonance is extraordinarily sharp. Thus oscillations of a frequency equal to that of the quartz resonator may be produced at will, the crystal playing the part of a tuning fork for high frequencies. Small bars or plates of quartz with a natural frequency of more than a million are easily produced, and the amplitude of the mechanical vibration may be so large as to fracture the crystal. The frequency at which resonance occurs may be determined with great precision by means of measurements of current or voltage in the tuned circuit to which the crystal is connected, but resonance may also be detected without the use of measuring instruments by means of the luminous resonator of Giebe and Scheibe[30, 31]. A small bar of quartz is supported between two electrodes, sealed into a glass bulb containing a mixture of neon and helium at a pressure of a few millimetres of mercury. At resonance the mechanical stresses in the crystal produce electrical changes on the surface of the quartz, and potential differences are set up, which are sufficient to produce luminous discharges in the gas mixture. Such a resonator forms a frequency standard which is extremely simple to use. It is merely necessary to connect the two electrodes to a coil of almost any size and to couple this magnetically to the source of oscillations to be investigated. The frequency is then slowly varied, until the glow appears, reaches maximum brightness and then disappears. By suitably adjusting the coupling, the frequency range over which a glow is seen may be made quite small, and by taking the resonant frequency as the middle point of this band an accuracy of setting of a few parts in a million may be obtained. The frequency of response is independent of the dimensions of the coil, since the coupling between the mechanical and electrical oscillations is so loose that there is no appreciable reaction. Resonators on this principle have been constructed for frequencies covering the range one thousand to ten million cycles per second. For the lower frequencies, 1000 to 20,000 c./sec. bars vibrating transversely are used; for the higher frequencies the bars are in longitudinal vibration. In order to excite the transverse vibrations, it is generally necessary to provide the rod with two pairs of electrodes. For example, if the four electrodes, two on either side of the bar, are

so connected as to produce an electric field along the electric axis, but in the opposite direction on the two sides, then at any instant one side of the bar will be shortened and the other lengthened, and therefore the bar is bent and transverse vibrations are produced. The temperature coefficient of frequency is of the order of 8 parts in a million per degree C. for the low-frequency resonators vibrating transversely, and about 2 parts in a million for the high-frequency resonators vibrating longitudinally. In the earlier forms of the instrument the crystal was loosely supported by the electrodes, but in the most recent form it is held in a definite position between the electrodes by a suitable metal support, to which it is tied by means of silk threads, the points of attachment being located at nodes of the mode of vibration to be used—generally an overtone. In this way the band of frequencies over which the resonator glows has been reduced to about 5 parts in a million, so that settings can be made to 1 part in a million. Such resonators have been found to remain constant with time to a few parts in a million.

Quartz oscillators. Although the resonator described above is the simplest form of frequency standard employing a piezo-electric crystal, it is generally more convenient for laboratory purposes to use a quartz oscillator, i.e. a quartz crystal maintained in continuous vibration by means of a thermionic valve. In the most convenient form the electrodes of the crystal are connected between the grid and cathode of a valve, and an oscillating circuit tuned approximately to the frequency of the quartz is included in the anode circuit. Over a considerable range of capacitance in the tuned circuit, electrical oscillations controlled in frequency by the vibrations of the crystal are produced. Crystals in the form of flat plates, thin rods, and rings have been used covering the frequency range 20,000 to 10,000,000 c./sec. It is to be expected that the natural frequency of a piece of quartz crystal will be constant to an exceptionally high order of accuracy, and that the quartz oscillator will form an excellent frequency standard. The natural frequency will, however, vary with the temperature and atmospheric pressure, and the frequency of the oscillator also depends to a comparatively small extent on the constants of the electrical circuit and the method of mounting the crystal, in particular the capacitance of the air gaps between the electrodes and the crystal. In the development of the quartz oscillator as a standard of frequency these factors have all been fully investigated. As regards the effect of temperature, it is a fortunate fact that the temperature coefficient of frequency for certain modes of vibration of quartz is positive while for others it is negative. In any actual vibrator there is nearly always a certain amount of natural mechanical coupling between two or more modes of vibration, and thus it is possible by a suitable choice of the dimensions of the piece of quartz to make the proportions of the various modes of vibration such as to make the temperature coefficient almost zero. Standard oscillators on this principle were first made by Marrison. Beginning with a flat circular disc, he found that the temperature coefficient was reduced by cutting a hole through the centre, and that for a hole of a certain diameter the temperature coefficient was less than 1 part in a million per degree C. The same principle has been employed more recently by Giebe and Scheibe in the standard quartz resonators and oscillators of the Physikalisch-

Technische Reichsanstalt. Bars of rectangular section vibrating longitudinally are used, and it is found that the temperature coefficient may be reduced by a suitable choice of the relative values of the length, width and thickness of the bar. The oscillator is in each case mounted in a double thermostat by means of which its temperature is maintained constant to about 0·001° C. so that changes due to temperature are made quite negligible. The capacitive coupling between the crystal and the electrodes must be kept constant, so that it is necessary to support the crystal in a definite position with respect to the electrodes, preferably midway between the various pairs, and in order to avoid friction at the points of support it is desirable that these shall be located at the nodes of the vibrator. Marrison's ring has a central nodal plane and it is suspended from a rod which passes through the hole in the disc, the sides of the hole being cut away so that line contact is made in this plane. The bar standards of the Physikalisch-Technische Reichsanstalt vibrate in overtones with two or more nodes, and are attached to supports at these nodes by means of silk threads for the lighter bars, or wire for the heavier ones. Lucas in this country uses a bar with a central node, and clamps the bar between the electrodes by two small pieces of cork at this point. Dye's ring oscillator has no nodes. It is supported by three loops of fine metal strip, such as is used for galvanometer suspension.

In order to determine the frequency stability of these oscillators, slow beats are obtained between two oscillators of nearly equal frequencies, and the time required to give a definite number of beats is measured. An accuracy of one part in 10^{10} may be obtained under suitable conditions, and it is found that oscillators of this type show relative frequency variations of the order of only $\pm 1 \times 10^{-9}$ over periods of an hour, and of $\pm 3 \times 10^{-9}$ over a period of 48 hours.

Tourmaline resonators and oscillators. Tourmaline also possesses piezo-electric properties and may therefore be used for the construction of resonators and oscillators. On the basis of its known mechanical properties it was anticipated that tourmaline would be more suitable than quartz for resonators of very high frequency, and a considerable amount of work has been done in this direction. Quartz discs a fraction of a millimetre thick often suffer from the disadvantage that they vibrate readily in different modes not far removed from one another in frequency, and change discontinuously from one frequency to another when the temperature or the circuit constants are varied. It was hoped that tourmaline would be better in this respect, but the results so far obtained suggest that the advantages of tourmaline are not very marked. Several investigators have successfully used tourmaline resonators at wave-lengths of the order of 1 metre, but the crystals (like quartz) are found to possess numerous inharmonic overtones[32]. The temperature coefficient of frequency varies from about -35 to -40 parts in a million per degree C.[33] and a logarithmic decrement of $1{\cdot}5 \times 10^{-4}$ has been obtained[34]. As a frequency standard for ultra-short waves tourmaline is most conveniently used as a simple resonator. The crystal is connected in parallel with the condenser of a tuned circuit, which includes a small glow-lamp. The frequency of the tuned circuit is approximately the same as that of the crystal. The tuned circuit is coupled to the oscillator to be measured, and the frequency of the oscillator is slowly varied. When it approaches

that of the tuned circuit the lamp glows, but when the actual frequency of the crystal is reached, and mechanical resonance occurs, the glow disappears. The accuracy of setting is about 2 parts in 100,000 at 15 Mc./sec.[35]

Multiplication and division of frequencies. Measurements of frequencies of almost any order of magnitude can now be made very simply in terms of a single frequency standard without any loss of accuracy by means of the processes known as frequency multiplication and division. Frequencies higher than the standard are compared directly with a suitable harmonic of the standard, and harmonics of a very high order may be obtained. The standard oscillator is coupled to another of approximately the same frequency but of a wave-form which is very rich in harmonics. The frequency is then held by the standard at a constant value. Any desired harmonic is obtained by coupling to the oscillator a circuit approximately tuned to the desired frequency, and amplifying the current induced in it. Harmonics up to the hundredth are readily obtained in this way, and it is then possible to repeat the process, using one of these harmonics as the fundamental frequency of the second stage and so on. In this way radio frequencies are measured in terms of a low-frequency standard, such as the tuning fork. The oscillator generally used for producing the harmonics is the multivibrator circuit of Abraham and Bloch.

Conversely, frequencies lower than the standard are obtained by taking an oscillating circuit such as the multivibrator, with a natural frequency which is approximately a submultiple, say one-tenth, of that of the standard, and coupling it to the standard oscillator. As a result of the non-linear characteristic of the low frequency oscillator it is held in step with the standard, and oscillations are obtained which are of a frequency exactly a submultiple of that of the standard. With two such stages a frequency of one-hundredth or even one-thousandth of that of the standard may be obtained, and in this way low frequencies may be measured in terms of a high frequency quartz oscillator. Any desired harmonic of the low frequencies thus produced may be obtained by frequency multiplication, and it will be realized that by a suitable combination of the two processes an accurately known frequency, quite close to any given value, may be obtained. The final comparison with the unknown frequency is made by the method of beats. Frequency division may also be carried out by means of a pentode[36] and frequency doubling by means of a triode[37].

The absolute measurement of frequency. In order to determine the frequency of a standard in terms of the fundamental unit of time, the mean solar second, it is necessary to link up observations of the vibrations of the standards with astronomical observations. Current of a frequency controlled by the standard oscillator is made to drive a small synchronous motor, which thus rotates at a speed which is an exact submultiple of the frequency of the standard. By means of a suitable train of gears the motor is made to close an electrical contact, at intervals of time containing a definite number of vibrations of the fork. This interval is frequently arranged to be very nearly 1 second, so that electrical impulses may then be obtained from the frequency standard, very similar to those obtained from the pendulum of a standard clock. These impulses are recorded on a chronograph, together with

impulses from a clock, or time signals from an observatory, and thus the mean frequency of the oscillator over various periods of time is determined.

The combination of standard oscillator, frequency-dividing circuit, and synchronous motor constitutes a time-measuring device, generally referred to as a vibration clock, and it is probable that the accuracy of such instruments exceeds that of any other form of clock. The "quartz clocks" of the Physikalisch-Technische Reichsanstalt have now been compared with signals from the various astronomical observatories over periods of several months. The mean rate of one clock over a period of six successive months showed deviations of only $\pm 0{\cdot}003$ sec. per day. Assuming that the rate of the quartz clock remains constant, the received time signals show discrepancies of a few hundredths of a second, and a systematic discrepancy has been revealed between time signals from Greenwich and Paris on the one hand and Potsdam and Hamburg on the other[38].

Chronographs. Several arrangements for automatically recording the performance of frequency standards or vibration clocks have been devised. By means of a rectifying circuit the slow beats between two oscillators of nearly the same frequency may be made to operate a relay, and thus an electrical impulse may be produced for each beat. These impulses and also those received from the clock pendulum may be counted by automatic counters, which print the results for given intervals of time on a paper band similar to that used in an adding machine. This arrangement was used by Marrison, and a similar one is used at the Physikalish-Technische Reichsanstalt. At the National Physical Laboratory one frequency standard is made to drive a chronograph drum at one revolution per second (by means of a synchronous motor and suitable gearing). Seconds impulses from the clock pendulum and other frequency standards are recorded on the drum, by a marking device which travels along the drum at 1 mm. per hour. Each marking circuit is automatically closed once per hour for a few seconds, and during this time the impulses from the clock are recorded as dots on the paper on the drum. The few successive impulses received on any one occasion make dots which are practically coincident, owing to the very slow motion of the marker, and thus the record consists of lines of dots, 1 mm. apart, made at intervals of 1 hour. The chart, which is about 18 cm. long, records a whole week's performance. If the rates of the various standards are the same, the lines are parallel to the axis of the drum. Since the speed of the paper is about 15 cm. per sec., time differences of a few thousandths of a second may be read, and thus the mean frequency determined with an accuracy which is limited only by the accuracy of the time signals available.

Marrison used an arrangement in which a circular transparent scale graduated with 100 divisions was driven at 10 revolutions per sec. by means of a synchronous motor controlled by one of the standard oscillators. The impulses to be recorded (e.g. beats between two standards) were made to operate a relay which caused a condenser to discharge through an induction coil. The secondary of the induction coil was thereby made to produce a spark across a gap near the revolving scale, and the portion of the scale illuminated by the spark was photographed on a slowly moving film. This arrangement records time differences to 0·001 sec.

A chronographic device with an extremely high speed of recording has recently been described by Kono[39]. A beam of light from an illuminated slit is reflected from the mirror of a vibration galvanometer on to a second slit. By means of suitable demultipliers, the frequency to be measured is reduced to a value f of about 500 or 1000 c./sec. Current of this frequency is supplied to the vibration galvanometer and, as the mirror vibrates, the light passing through the second slit is interrupted at the frequency of $2f$. Seconds signals are obtained from the pendulum of a standard clock by photoelectric means, and are supplied to a second vibration galvanometer, and the interrupted beam of light is reflected from the mirror of this second galvanometer and then focused on to a film carried by a drum which rotates slowly. Each impulse received from the clock gives a sudden deflection to the interrupted beam, and causes the spot of light to move rapidly to and fro across the film. The record therefore consists of rows of dots, the dots being controlled by the frequency f, and the rows by the clock pendulum. The frequency f is deduced from the relative phase displacements of the dots in successive series of rows. The speed of the spot of light on the drum was 20 metres per second, and it is claimed that the timing error is only 35 microseconds, so that frequency may be measured with an accuracy of better than 1 in a million, over a time interval of only 30 seconds.

§ 5. THE ABSOLUTE MEASUREMENT OF CAPACITANCE

The Maxwell bridge. Maxwell's bridge for the measurement of capacitance consists of a simple Wheatstone bridge network in which one resistance arm is replaced by the condenser to be measured and some form of commutator arranged to charge and discharge the condenser at a constant rate. This combination is equivalent to a resistance, since it transmits a definite mean current for a given applied voltage and, provided the galvanometer indicates mean current, the bridge may be balanced and capacitance determined in terms of the resistances in the network and the frequency of the charging of the condenser. Thus capacitance is determined in terms of resistance and time, and the method becomes an absolute one, if the resistances are determined by an absolute method. It has probably been used more widely than any other method for the determination of the capacitance of the fundamental standard condensers now in use in various parts of the world. The most recent investigation of the method was carried out at the Electrotechnical Laboratory, Tokyo[40], but it follows so closely one carried out by Curtis and Moon[41] at the Bureau of Standards in 1927, that the two should be considered together. The method was previously used at the Physikalisch-Technische Reichsanstalt by Gruneisen and Giebe[42], and there is little doubt that an accuracy of 2 or 3 parts in 100,000 was obtained in each of these investigations. The method is only applicable with high accuracy to condensers which are sensibly perfect, i.e. to standard air condensers of negligible leakage resistance, and showing no appreciable amount of absorption. As a check on the perfection of the condenser it is necessary to make measurements with widely different rates of charging, e.g. Curtis and Moon used rates of 100 and 1000 per sec. The bridge equations used for these precision

measurements are identical with those given by J. J. Thomson in 1883, although Curtis and Moon have considered the theory in somewhat greater detail than was originally given. For example, when the bridge is balanced the galvanometer is not absolutely stationary but vibrates about the zero position, since the steady current passing through the resistance is balanced by a series of impulses arising from the condenser charges. Curtis and Moon introduced into their equations a term representing this motion. The final equation showing the condition under which the time integral of the current through the galvanometer is zero is fortunately not affected, but the question as to how the observer judges whether the time integral of current is zero requires consideration. The spot of light from the galvanometer shows, not a sharp line, but a band sharper on one edge than the other, and it is shown that this should be set so that the zero of the scale is inside the band and at a distance of one-third of its width from the sharper edge. It is essential to the method that the galvanometer shall integrate the current correctly, but Curtis and Moon found that many galvanometers do not do so, and the experiments of Yoneda and Yanaguchi supply additional evidence of this. Galvanometer coils wound on aluminium formers were invariably unsatisfactory, the effect being probably due to slight ferro-magnetic properties in the material. The current impulses produced in the galvanometer at the instant of charge are large compared with mean value, and these impulses may leave traces of residual magnetism in the former, which will produce a shift in the zero of the galvanometer. The effect of the galvanometer current on the magnetic field in which it swings also requires consideration, and it is shown that this also causes incorrect integration, or a change of zero if the coil is not placed symmetrically in the field of the permanent magnet. The symmetrical position can be found experimentally, but for measurements on large condensers Curtis and Moon found it advisable to use a special permanent magnet such that the field it produced was practically unaffected by the current in the coil. It was of cobalt steel, large in dimensions, magnetized first to saturation and then partly demagnetized by an alternating current. The material was therefore in a very hard magnetic condition. The various investigations have differed very considerably in the means adopted for charging and discharging the condenser at a constant rate. At Tokyo a commutator was driven by a synchronous motor, the speed of which was controlled by the frequency of a valve oscillator, which supplied the current to its rotor. A small d.-c. motor was also directly coupled to the synchronous motor in order to supplement its power. A stroboscopic disc on the shaft was illuminated by a neon lamp energized by the valve oscillator, and served to indicate synchronism. At the Bureau of Standards a tuning-fork interrupter with contacts of platinum dipping into mercury cups was used. The fork was of elinvar (giving a very low temperature coefficient of frequency) and it was maintained in vibration by a valve circuit. These devices give speeds which are constant to a very high order. The tuning fork is probably superior to any other form of this device so far produced but it could not be successfully used for frequencies much greater than 100 per sec., and for higher frequencies a commutator was also used at the Bureau of Standards. The frequency of the commutator was measured in terms of signals from a standard

pendulum or chronometer by means of a chronograph, but the frequency of the fork was measured by a stroboscopic device. The prongs of the fork carried vanes perforated by slits forming a shutter which opened and closed twice per cycle. This shutter was illuminated and an image of the slit projected on to a screen by means of a small concave mirror carried by a free pendulum. When the fork and pendulum were in synchronism a series of images of the slit formed on the screen appeared to be stationary, and observations of their motion determined the frequency. An accuracy of 1 part in 10^6 could be obtained by observations extending over four minutes. A quarter-second pendulum made from a single piece of fused quartz, and mounted in an air-tight case was used, so that its temperature and pressure corrections were negligible. The errors in the measurements of time required for Maxwell's method may now be regarded as quite negligible, and observations are reproducible to better than 1 part in 100,000, but some doubt remains as to whether the ideal electrical conditions of theory can in practice be satisfied to an extent that will give an accuracy better than 2 or 3 parts in 100,000. Higher accuracy than this would be of no practical importance for measurements of capacitance alone, but as such measurements form one step in an important method for determination of resistance in absolute measure, the question is of some significance.

The Carey Foster bridge. At the National Physical Laboratory basic measurements of capacitance are made by means of the well-known Carey Foster bridge for the evaluation of capacitance in terms of resistance and mutual inductance. The method has the advantage that it does not demand the use of perfect condensers, and that it may be used with alternating current of any frequency up to about 2000 c./sec., but as the frequency is increased, the corrections necessary on account of capacitance effects in the coils of the mutual-inductance standard become a serious limitation. Details of the corrections have been given in a recent paper [43, 44], which also gives an account of comparative measurements made on standard mica condensers, first at the Bureau of Standards, where the condensers were compared with standard air condensers calibrated by means of the Maxwell bridge, and afterwards at the National Physical Laboratory, where the measurements were made by means of the Carey Foster bridge in terms of a primary standard of mutual inductance. The results agreed to about 1 part in 10,000, but a closer comparison could not be obtained owing to variations in capacitance of the mica condensers. The accuracy obtainable by means of the Carey Foster bridge is greatest at the lower frequencies (of the order of 100 c./sec.) where the corrections due to capacitance in the inductive coils and residual inductance in the resistance coils become negligible, and at these frequencies an accuracy of perhaps 2 parts in 100,000 may be obtained. There is, however, as yet little point in obtaining the highest possible accuracy since standard condensers with a stability of the required order have not yet been produced.

Standard air condensers. Maxwell's bridge demands the use of perfect condensers, and therefore laboratories using this bridge as their basis of capacitance measurements have made investigations of standard condensers in considerable detail. Mica condensers, although satisfactory for most ordinary purposes, such as measurements of capacitance by ballistic and alternating current methods, are

unsatisfactory for the Maxwell bridge and electrometer methods, since the mica always shows appreciable absorption. Thus the capacitance may vary with frequency by a few parts in ten thousand over the audio-frequency range, while the value given by ballistic galvanometer methods will differ from the a.-c. value by a few parts in a thousand, for the best mica condensers obtainable. Air condensers are therefore necessary for use as standards. The standard air condensers of the Physikalisch-Technische Reichsanstalt have recently been described in considerable detail by Giebe and Zickner[45]. Early measurements showed that condensers built up with flat metal plates were more stable than those with plates in the form of concentric cylinders. The flat plate form is also the more compact, and has therefore been used for all modern condensers. It will be realized that in order to get capacitances of the order 0·1 μF. it is necessary to use very many plates with a small air gap between them, and the problem of providing perfect insulation between these plate systems, and yet making the whole structure so rigid that no appreciable changes of capacitance occur, is a difficult one. The thickness of the metal plates, their tendency to sag or warp, the total weight of each plate system, and the size of the air-gap are factors which are all closely connected with each other and with the stability of the capacitance. The nature and form of the insulating material is also an important factor. Each plate system is built up on a stout base plate, with supporting pillars and separating washers locating the plates in the ordinary way. The two base plates are separated from one another by small pillars of insulating material. They are then firmly clamped to the base of the condenser by a stout clamping ring and bolts, additional insulating pillars separating the base plates from the final base and the clamping ring. The insulating material is thus all in compression. Transparent fused quartz is used for condensers intended for a.-c. work since this gives the lowest power factor, although reconstructed amber is found to give a higher leakage resistance and therefore is better for d.-c. measurements. In general the plates are of rolled aluminium sheet about 1 mm. thick, and the air gap is about 2 mm., but certain of the larger condensers have plates turned from castings of copper-aluminium. These plates are 3 mm. thick, and it is found to be possible to use them with an air-gap of only 1 mm. The plates and base of each system must of course be bored with holes through which the pillars supporting the other system may pass. The number of supporting pillars is increased, as the size of the plates is increased, in order to retain the necessary rigidity. The largest condensers (0·1 μF.) have plates of diameter 35 cm., each system being supported by five pillars, four distributed round the circumference and one near the middle. In the first few years after their manufacture the large condensers are found to increase in capacitance by perhaps 2 or 3 parts in 1000, but ultimately a more stable condition is reached and the stability is then of the order of ± 2 parts in 10,000 over a period of several years. The smaller condensers (less than 1000 $\mu\mu$F.) show a stability of this order from the time of their manufacture. Attempts to improve the stability by annealing the plate systems have been made, but so far the evidence available is insufficient to show any definite result. If the temperature of the condenser is maintained constant, the stability over a period of several hours or even days is better than 1 part in 100,000.

The temperature coefficient of capacitance is of the same order as the coefficient of linear expansion of the metals forming the plate systems, viz. 2 or 3 parts in 100,000 per 1° C., but for temperature changes of 5° or more the changes of capacitance are not in general reversible. Generally speaking, condensers are much less satisfactory as permanent standards than inductances or resistances, and although small condensers can be made much more permanent than large ones, they cannot usually be measured with so high a degree of accuracy.

The capacitance of two condenser units in parallel is only strictly equal to the sum of the capacitances of the two units if each unit is completely screened, and the leads connecting them have zero capacitance. These conditions are realized in the Physikalisch-Technische Reichsanstalt condensers by providing the plate systems with sockets, flush with the outer case of the condenser, instead of terminals. The outer cases are completely closed except for small holes through which the sockets are accessible. Condensers are connected in parallel by standing one on the top of the other. The sockets in the base of the upper condenser register with those in the top of the lower one, and they are connected by small plugs, which are completely enclosed by the two sockets and therefore make no contribution to the capacitance. Variable condenser units may be added in the same way and thus any desired capacitance built up. A base plate provided with similar sockets and suitably connected to the measuring apparatus must also be used with every combination, but the capacitance of this plate and connecting leads becomes the zero of capacitance of the system, and is allowed for by a zero reading in the usual way.

Standards of calculable capacitance. Although the formulae for the calculation of capacitance are comparatively simple, calculable standards of capacitance are little used, most capacitance measurements being made in terms of absolute standards of inductance. The reason for this is that it is very difficult to produce a condenser whose linear dimensions may be determined with the highest precision and whose capacitance is at the same time large enough to be generally useful. Standard condensers in the form of concentric cylinders with a capacitance of the order of 50 to 100 $\mu\mu$F. have been made and are frequently used for high voltage work. Churcher and Dannatt[46] have described such a standard in which an accuracy of a few parts in 10,000 was obtained. This condenser was designed for a maximum working voltage of 300 kV., the cylinders having diameters of about 40 and 20 cm. The accuracy was checked by a measurement of capacitance in terms of mutual inductance and resistance standards, using for this purpose a modification of the Carey Foster bridge. In order that the simple formula for a cylindrical condenser may be applied, it is necessary to eliminate end effects, i.e. to measure the capacitance of a definite length near the middle of the cylinders, excluding the end portions where the field differs appreciably from that for infinite cylinders. This is done by an application of the principle of the guard-ring. The capacitances of the connecting leads are of course of great importance in dealing with such small values of capacitance, but these may be eliminated either by the use of suitable shielding, or by measuring them separately when taking the zero reading.

Condensers of very low value—micro-condensers. In the rare cases in which

capacitance standards of 1 $\mu\mu$F. or less are required, absolute standards become almost essential. H. Lampe[47] has described the use of a cylindrical arrangement for such values. A fixed outer cylinder is used, and wires of known length and diameter are screwed in turn into an insulated and screened socket, which is located at one end of the cylinder on its axis. Thus condensers of the order of 0·5 $\mu\mu$F. are obtained.

Variable air condensers capable of being read to 0·001 $\mu\mu$F. are now fairly extensively used for such work as the measurement of the inter-electrode capacitance of thermionic valves. They usually take the form of concentric cylinders, one of which is capable of axial adjustment by means of a micrometer screw. J. Clay[48] has recently described such an instrument made in steel with great precision to serve as an absolute standard. End effects are of course very considerable in these instruments, but with a suitable design they are sensibly constant over a considerable portion of the range, and once the extent of this portion has been found experimentally, it may be used as an absolute standard for differences of capacitance. A somewhat similar condenser has been described by Bartlett[49]. Clay found that the middle portion of the scale of his condenser was uniform to 1 part in 10,000 and claimed that by its use capacitance of the order of 1 $\mu\mu$F. could be measured with an accuracy of this order.

The electrometer method of capacitance measurement. Both Clay and Lampe used the electrometer method for the measurement of capacitance in terms of their absolute standards. The two condensers form two arms of a Wheatstone bridge network. Two resistances of the order of 100,000 Ω. form the other arms. A battery of say 50 volts is connected across the two resistance arms and a sensitive electrometer occupies the place of the usual galvanometer. Lampe claims a sensitivity of 1 part in a million for this method, but it is only applicable to condensers showing no leakage conductance, and complete shielding is of course essential. He compared it with a resonance method using current of radio frequency and found that the two methods gave results identical to 1 part in 1000, which was the order of error of the measurement at radio frequency.

§ 6. THE MEASUREMENT OF DIELECTRIC CONSTANT, CAPACITANCE, AND CONDUCTANCE AT HIGH FREQUENCIES

Methods involving the use of stationary waves. In the last few years there have been several developments of the old methods introduced by Drude for the measurement of dielectric constant and conductance (absorption coefficients) at frequencies of the order 10^8 c./sec., or wave-lengths of the order of a metre. The Lecher system of two straight parallel wires is employed and it is coupled to an oscillator tuned to the required frequency, so that electromagnetic waves of the same frequency travel down the wires, and suffer reflection at the far end, which may be either open or closed. The length of the parallel wires is adjustable, and when it is an integral number of half wave-lengths resonance is established, and a series of nodes and antinodes distributed along the wires may be found by means of a suitable detector

of oscillations. In order to measure the dielectric constants of liquids, Drude placed the wires in a long trough, which could be filled with the liquid. He then determined the wave-length in air and in the liquid, and from the ratio obtained the refractive index n of the liquid. For liquids in which the conductance is negligible the dielectric constant is equal to n^2. Drude used damped trains of waves obtained from a spark generator, and as detector he employed a neon discharge tube, which spanned the parallel wires and glowed with maximum brightness at the antinodes. The effective length of the wires was varied by moving a conducting bridge along them. It is remarkable that with such comparatively crude apparatus results which are probably correct to about 2 per cent were obtained. The method was used for a few standardizing liquids which were available in sufficient quantities to fill the trough. Liquids available in only small quantities were investigated by a second method, in which the distant ends of the parallel wires were closed by a small test condenser in a glass bulb which could be filled with the liquid. The length of the wires was adjusted by a telescopic tube arrangement, and the distance between the condenser and the node nearest to it was measured. The position of the node was accurately found by moving a conducting bridge along the wires until resonance occurred. This distance varies with the capacity of the condenser, i.e. a given capacitance may be regarded as equivalent to a certain length of the parallel wire system, and, by filling the condenser with liquids standardized by the first method, the apparatus could be calibrated in terms of dielectric constant, and other liquids then measured. When energy was absorbed by the liquid in the test condenser owing to its conductance, the glow of the discharge tube was less intense, and a measure of the conductance was obtained by comparing the intensity of the glow in any given case with that obtained when an electrolytic solution of known conductivity was used.

Drude's first method was modified in 1930 by Drake, Pierce, and Dow[50] into a form capable of high precision. The two parallel wires were replaced by a single wire and a metal tube surrounding it, and capable of holding the liquid under examination. In this way the electric field was confined to the liquid, and the influence of surrounding objects was greatly reduced. Continuous oscillations produced by means of thermionic valve oscillators were employed. The oscillator was completely enclosed in a shielding box in metallic connection with one end of the outer tube, and the corresponding end of the central wire was coupled to the grid of the valve-oscillator by means of a small variable condenser. Resonance was detected by observing the anode current of the oscillator. As long as the wire system is out of tune, the anode current remains constant, but when resonance occurs an increased load is suddenly thrown on to the oscillator, and there is a marked change in the value of the current. The tuning was carried out by moving along the tube a brass plunger which made contact between wire and tube at a definite position along the axis. The total length of the concentric system was 3·7 m. and with water in the tube settings could be made with an accuracy of about 0·1 mm. The theory of the system considered as a transmission line has been worked out, including the effects of resistance and inductance of the concentric conductors, and of conductance

as well as capacitance in the separating medium. The dielectric constant of distilled water was measured over the temperature range 3° C. to 60° C. and the frequency range 12 to 76 Mc./sec. The value was $78{\cdot}57 \pm 0{\cdot}05$ at 25° C. and was independent of the frequency over this range. The effect of increasing the conductivity of the liquid by adding KCl (up to $\frac{1}{70}$ normal) was to diminish the wave-length in the liquid considerably, but when allowance was made for the effects of conductance the values obtained for the dielectric constant were within 1 per cent of that for pure water. The accuracy of the method diminishes as the attenuation due to conductance increases. Malone, Case and Ferguson[51] have also recently attempted to apply Drude's first method to conducting solutions, but with increasing conductivity they obtained different results with different systems of resonance detection. In one series of experiments their wires were closed through a thermomilliameter at the oscillator end, so that the current in wires at this point was read directly, and a maximum value observed at resonance. In other experiments a valve voltmeter was used to measure the voltage across the wires and the maximum voltage observed. Since current resonance is observed in one case and voltage resonance in the other it is not improbable that mathematical analysis will account for the discrepancy.

M. Seeberger[52] following K. Kohl[53] has developed a modification of Drude's methods and has used it for the measurement of the dielectric constant and conductance of several liquids over the frequency range 1200 to 2400 Mc./sec., i.e. wave-lengths in air of 12 to 24 cm. Two vertical parallel wires were used, their lower ends being coupled to the valve oscillator and their upper ends passing through the flat bottom of a glass vessel containing the liquid. The ends of these wires were closed by a metal reflecting plate, parallel to the bottom of the vessel, and the thickness of the layer of liquid traversed by the waves was adjusted by raising or lowering this plate. In addition the total length of the wires was adjustable. The detector used was a very small crystal rectifier connected to a galvanometer in an untuned circuit, so that it would cause very little disturbance of the field. Starting with the reflecting plate very near the bottom of the vessel, the length of the wires was adjusted for tuning, and the detector set at an antinode. The reflecting plate was then moved slowly upwards so as to increase the thickness of liquid traversed, corresponding readings of layer-thickness and detector-reading being noted. As the layer of liquid became thicker the detector reading passed alternately through maxima and minima, but the amplitude of the alternations of reading diminished, owing to the absorption in the liquid. The wave-length in the liquid was determined by the distance between successive maxima, while the absorption index was calculated from the decrement of the alternations.

At these very high frequencies it is also possible to use free waves in a liquid. Thus a valve generator and metal reflecting plate were completely immersed in a liquid. As the distance between the two increased, the anode current of the valve generator varied periodically, indicating the existence of stationary waves as in the experiments of Drake, Pierce and Dow. The distance between successive nodes can thus be measured in air and the liquid, and the refractive index, determined. The absorption index was also directly measured by the use of free waves by placing

oscillator and detector on either side of a layer of liquid, and observing the diminution of detector-reading as the depth of the liquid increased. Several liquids, mainly mixtures of various alcohols and water, were examined and anomalous dispersion found in some cases, e.g. for water between wave-lengths of 12·6 and 19 cm. A method in which the parallel wire system passed through a small cell containing a layer of liquid, this cell being at the middle of the wires, was found to give doubtful results. The difficulties of such arrangements are largely due to the fact that reflections occur at all discontinuities in the path of the waves, i.e. at each surface of the cell, and at any point on the wires at which sliding contacts are made.

Comparative measurements of conductivity at very high and low frequencies have also been made by Glowatzki[54] following Zahn and Rieckhoff. At the high frequencies (3×10^8 c./sec.) a Lecher parallel wire system was used, the distant end being closed by a circular loop of wire. The liquid to be examined was contained in a cylindrical vessel which was placed in the centre of this loop. The eddy currents generated in the liquid damped the oscillations by an amount depending on the conductivity of the liquid. The position of the vessel was accurately defined by means of a geometrical mounting. The detector was a barretter arrangement; four very thin platinum wires arranged in the form of a Wheatstone bridge, supplied with direct current and balanced in the usual way. The high-frequency current passed through one of the wires, raised its temperature, and therefore its resistance, and so caused the galvanometer to deflect. An extension of the cylindrical vessel, also filled with the liquid, was provided with internal electrodes and these were used for the measurement of the conductivity at low frequencies by the ordinary bridge method.

Resonance methods. The most widely used method of measurement of capacitance at radio frequencies is that of substitution in a circuit tuned to resonance with an oscillator to which it is loosely coupled. The unknown condenser and a standard variable air condenser are connected in parallel with an inductive coil. The standard condenser is set at a point low on its scale and the circuit tuned to resonance. The unknown condenser is then disconnected, and the circuit again tuned by varying the standard condenser. The difference between the two readings of the standard gives the value of the unknown condenser. Many investigations, chiefly measurements of dielectric constant, have recently been made by this method, but for the most part they call for no special mention. As in all measurements of capacitance, it is necessary to ensure that the disconnection of the unknown condenser does not itself introduce appreciable capacitance into the circuit. An arrangement devised with this object is described by Sugden[55]. Again it is important that the frequency of the oscillator shall remain constant during the experiment; oscillators controlled in frequency by a piezo-electric crystal are now widely used. Several methods of detection of resonance are available; the current in the circuit may be measured by means of a thermal instrument, and the circuit adjusted for maximum current, or the voltage across the circuit may be measured by means of a thermionic voltmeter (or any instrument of high impedance) and the adjustment made for maximum voltage. In high precision work it may be important to recognize that the condition

of maximum current (in the circuit external to the coil) is not quite the same as that of maximum voltage across the coil, owing to the fact that the coil itself possesses capacitance as well as inductance, and the current passing through the capacitance of the coil is ignored by the ammeter. When measurements are required on condensers possessing appreciable conductance, the voltage method of detection possesses the important advantage that the point of resonance is unaffected by the conductance, which becomes merely a shunt across the voltmeter. Also the resonance curve connecting voltage and condenser reading is symmetrical, so that the condenser reading corresponding to the maximum may be accurately found by taking the mean value of any two readings corresponding to the same voltage reading. Another method of resonance detection requires no voltmeter or ammeter, but depends on the fact that when a circuit is in resonance it possesses zero reactance. If therefore such a circuit is coupled to the tuned circuit of an oscillator, the frequency of which is determined by the capacitance and inductance in this tuned circuit, the frequency of the oscillator will be unaffected by the coupled circuit so long as it is in resonance, but if the condition of resonance is departed from, positive or negative reactance is coupled to the oscillator circuit and its frequency goes down or up as the case may be. If a standard oscillator of constant frequency is available, such small changes of frequency may be detected by the method of beats, and thus the point of resonance found with great sensitivity. The sensitivity may be increased by adjusting the two radio-frequency oscillators so that they produce an audible beat note, and comparing this note with one of nearly the same frequency obtained from a standard audio-frequency oscillator or tuning fork. This so-called "double beat method" is probably the most sensitive indicator of resonance available. The method of observation is to note the changes of frequency occurring when the test circuit is opened and closed, and to adjust the capacitance of this circuit until the changes are reduced to zero. Full practical details have been given by Colebrook[56].

Resonance methods of measurement of resistance and conductance are of two kinds, which may be called resistance-variation and reactance-variation respectively. In the resistance-variation method the circuit is tuned to resonance and the maximum current or voltage noted. A known resistance is then added to the circuit and the maximum reading again noted, the induced e.m.f. being maintained constant. Since the tuned circuit possesses zero reactance, the current in each case is proportional to the total resistance, and thus the total circuit resistance is readily calculated. The condenser or coil whose resistance is to be determined is then added to the circuit, and the whole process repeated. The difference between the two values for the circuit resistance determines the resistance or conductance of the added component.

In the reactance-variation method the maximum current or voltage at resonance is first noted as before. The impedance of the circuit is then changed by adding a known reactance and the current again noted. This may be done by altering the capacitance of the tuning condenser by a known amount. The extent of the reduction of current or voltage produced by a known change of reactance enables the resistance or conductance of the circuit to be calculated. In practice the detector

usually gives a deflection proportional to the square of the current or voltage, the value at resonance is noted, and then the capacitance is varied until the deflection is halved. Two such capacitance readings C_1 and C_2 are obtained, one greater and one less than the value C_r at resonance, and the total conductance G of the circuit is simply given by

$$G = \tfrac{1}{2}(C_1 - C_2)\,\omega,$$

ω being the angular frequency. In order to measure the conductance of any given component, this is added to the circuit and the process repeated, the required value being obtained by difference. The known change of reactance may also be produced by a small change of frequency, the capacitance and inductance remaining unaltered, and the conductance calculated from frequency observations. This procedure is very useful when no variable inductance or capacitance is permissible. These methods have recently been used by Rohde and Schlegelmilch[57] in an investigation of the dielectric constant and conductance (expressed as phase angle or power factor) of various solid dielectrics at frequencies up to 50 Mc./sec. (wave-length 6 m.). For wave-lengths down to 20 m., the resistance-variation method could be employed without any special difficulty. At higher frequencies, the inductance coil consisted of a single turn of thick copper rod, and the condenser of a pair of metal discs between which the specimen was placed. Thus it was necessary to use the frequency variation method of conductance measurement. It is stated that the ordinary thermionic voltmeter is not satisfactory at such high frequencies and the detector used was one of special form, comprising a rectifying diode of very small inter-electrode capacitance, with an almost point filament, and a string electrometer for observing the rectified voltage. The conductance measurements were obtained from the difference between the readings when the sample was inserted, and when there was only air between the metal plates. The gap between the plates was adjustable and the capacitance was adjusted to be the same in the two cases, sample in and sample out. This makes the conditions of the two sets of measurements almost identical, so that although there are likely to be considerable departures from the conditions assumed in simple theory (there is no screening for example) most of the errors are likely to cancel out on taking the difference. It is important that the same condenser plates are used for the air condenser, and that containing the dielectric, since at such high frequencies any ordinary standard air condenser would possess an appreciable conductance, owing to the dissipation of power in the plate systems on account of their resistance. For a given resistance of the plate systems the equivalent conductance is proportional to the square of the capacitance and the square of the frequency—so that it is only by the use of the same capacitance, and the same plates, that the conductance of the air condenser could be eliminated. It is to be noted that ranges of frequency covered by the resonance methods and the stationary wave methods may be considered to overlap, so that some check on their accuracy should be possible. The stationary wave method has however not yet been successfully applied to solid dielectrics in sheet form, such as glass, amber, quartz and the like, and it is such materials that Rohde and Schlegelmilch were mainly concerned with.

High-frequency oscillators. It is to be noted that the improvements in Drude's methods for the measurement of dielectric constants at high frequencies have only been made possible by the invention of sources of continuous oscillations at such frequencies. All recent workers in this field have used thermionic valve oscillators. Ordinary valve oscillators, of which the frequency is mainly determined by that of a tuned circuit forming part of the network, are of two types; the reaction type, in which the production of oscillations depends on the coupling of the anode and grid circuits; and the dynatron type, which depends on the existence of a so-called negative resistance in the valve, i.e. over the operating range of electrode potentials the current must increase as the potential falls. The frequency of such oscillators may be varied at will, simply by changing the inductance and capacitance in the tuned circuit, and frequencies from very low values up to 50 Mc./sec. (wave-length 6 m.) may readily be produced. At still higher frequencies, however, the mechanism of production of such oscillations fails; inertia of the electrons becomes an important factor in determining their motion; and with certain valves it is found to be possible to maintain continuous oscillations in virtue of the electron inertia of the system. The range of frequency that can be produced is somewhat limited, since it depends critically on the time of transit of the electrons from cathode to anode; the frequency is controlled by the properties of the valve itself rather than those of the associated circuits. The earliest form of electronic oscillator and the one which has been used for most of the work previously referred to, is that of Barkhausen and Kurz[58]. The grid of a triode is raised to a high positive potential, and the anode is made slightly negative, with respect to the filament. It is then found that very high-frequency oscillations can be maintained in a circuit connected between grid and anode, grid and filament, or anode and filament. These oscillations depend on the periodic motion of electrons in the valve, and are only maintained when the resonant frequency of the circuit bears a certain relation to the natural frequency of the electron system. The explanation usually given is that electrons are accelerated in passing from filament to the grid; some of these pass through the grid, are retarded in the grid-anode space, brought to rest near the anode, once more attracted towards the grid, through which some again pass, and so on. It will be evident that the frequency will depend on the electrode distances and the voltages. A summary of recent work on such electronic oscillations has recently been given by Megaw[59].

The efficiency of the Barkhausen-Kurz type of oscillator is comparatively small, perhaps 1 per cent, but Kozanowski[60] has obtained an output of 5 watts at a wave-length of about 70 cm. by an arrangement of two transmitting valves. A considerably higher efficiency may be obtained from the magnetron oscillator, and it is probable that this will become widely used in high-frequency measurements. The properties of an efficient type now generally available have been described by Megaw[61]. It consists of a diode with a straight filament and a cylindrical anode divided lengthwise into halves, each being provided with an outgoing lead. The valve is mounted between the poles of a magnet, which provides an axial uniform magnetic field in the space between the electrodes. Thus electrons leaving the filament and proceeding towards the anode are deflected by the magnetic field, and

when the field-strength passes a certain limiting value they travel round in closed curves and do not reach the anode. When this condition is reached the mean anode current suddenly falls almost to zero, and it is then possible to obtain powerful high-frequency oscillations in a circuit connected to the two parts of the anode, the oscillations in the two being in opposite phase. Oscillations of two types may be produced. Within a certain range of anode voltage, a decrease of voltage on one anode is found to be accompanied by an increase of current, i.e. the magnetic field introduces a negative resistance into the valve characteristics, and if the valve is coupled to a tuned circuit, oscillations are generated at a frequency determined by that of the circuit. Thus any desired frequency may be produced provided the periodic time of the oscillation is large compared with the time of transit of the electrons from filament to anode. At very high frequencies when this condition no longer holds, the negative-resistance property is no longer effective, and oscillations cease, as in the case of the ordinary reaction or dynatron oscillator. The magnetron has, however, advantages at these very high frequencies in several respects; its internal inductance and capacitance can be made very small, so that the upper limiting frequency is determined solely by the effects of electron inertia; also as the limiting frequency is approached, the output does not steadily diminish as does that of the ordinary triode, but the first effect of electron inertia is to increase the output, decreasing it again at still higher frequencies. Thus certain magnetron valves are specially suitable for producing oscillations of wave-length 1 to 5 m.

A second type of oscillation may be obtained at frequencies higher than the upper limit described above. For these, the magnetic field-strength must be near the critical value at which electrons just fail to reach the anode. The mechanism of the oscillations is not completely understood, but their periodic time appears to be twice the time of transit of the electrons from filament to anode, and they are ascribed to the motion of a certain proportion of the electrons in circular orbits. Oscillations of wave-length less than 50 cm. are readily produced in this way, and Megaw has obtained an output of 0·5 watt at a wave-length of 30 cm. with a small valve.

§ 7. HIGH-FREQUENCY BRIDGE MEASUREMENTS

For the measurement of impedance, including resistance, inductance and capacitance, bridge methods have great advantages over other methods, since the readings are unaffected by fluctuations of the current, and the accuracy is not limited to that with which deflections can be read. Several attempts have therefore been made to apply bridge methods to measurements with currents of radio-frequency. All the well-known bridge circuits for the measurement of capacitance, inductance, phase angle, resistance, and frequency are theoretically applicable to current of any frequency, but it is a matter of common experience that a circuit which gives satisfactory results at 100 c./sec. may give appreciable errors at 1000 c./sec., very large errors at 10,000 c./sec. and be quite unworkable at higher frequencies, although a suitable generator and detector of balance may be available. The difficulties arise from the fact that as the frequency is raised, the currents are

no longer confined to the meshes of the bridge network, but flow through space between any two components of the circuit which are not at the same potential. The displacement current between any two conductors in the system is proportional to the frequency, to the capacitance between them, and to their potential difference, and at high frequencies such currents become comparable with those circulating in the conductors of the network, so that the network postulated by theory is no longer realized. The currents may, however, be confined to definite paths by suitable screening, and in this way bridge circuits have been built so as to give satisfactory results at frequencies up to a million. The displacement currents flow not only between the various components of the bridge arms, but also between each component and other neighbouring objects including the observer, the walls of the room, etc. Thus all earth-connected objects in the neighbourhood are of necessity included in the measuring circuit, and due attention must be paid to their position in the system, as the currents flowing to earth are apt to be large.

The general principles governing the screening are as follows: (*a*) Each arm of the bridge should be completely enclosed by a conducting screen, which is usually connected to one terminal of the unit, the other (here called for convenience the insulated terminal) being accessible through a small hole in the screen. This construction ensures that all the current entering the arm by the insulated terminal must leave it by the screen terminal. (*b*) In connecting together the four arms of the bridge, the insulated terminal of one arm should be connected to the insulated terminal of the next, and the screen terminal should be connected to the screen of the arm on the other side of it. Thus all the screens become connected to two opposite corners of the bridge, the other two opposite corners being occupied by insulated terminals only. Since the screens completely enclose the four arms, and they are all connected to two points only, the only displacement or capacitance currents linking the four arms are those between these two points, and if these two points are those to which the current generator is connected, such displacement currents merely represent a shunt across the generator and do not affect the conditions of balance of the four arms of the bridge. (*c*) The current generator and the detector of balance must each be completely enclosed in a metal screen, which should be connected to earth. Two connecting leads must of course be brought out from the shield of each instrument for connexion to the appropriate points of the network, but these leads must also be enclosed in metal tubes which are earth connected. (*d*) The assembly comprising the four arms of the bridge is also completely enclosed by an earth-connected shield, so that no current passes from the measuring system to the body of the observer, and movements of the observer have no effect on the current distribution. (*e*) Capacitance currents from the screens of the four arms to the earth-connected shield must now be considered. Frequently one of the points to which these screens are connected may be connected to the earthed screen, and then the capacitance between the other point and earth becomes merely a shunt across the current-generator and does not affect the balance of the bridge. (*f*) The effects of the capacitance to earth of the detector of balance also require consideration, and this is usually the chief difficulty. It is essential that the

only current passing through the detector shall be that received from the appropriate two corners of the bridge, but it is obvious that current may also flow through the detector via capacitance to earth at various points of the complete system of bridge, generator and detector, and earth. The difficulty has been overcome in two ways. The simplest[63, 65, 66, 67] consists of using only bridges with equal ratio arms, and then arranging the detector so that its capacitance to earth is balanced with respect to the two points at which it is connected to the bridge; either by the use of a symmetrically wound and screened transformer for coupling the detector to the bridge, or else by the equalization of the capacitance by means of variable condensers. The result is that when balance is obtained, capacitance current received by the detector from one pair of arms is balanced by an equal current from the other pair and no false balance is obtained. A more general solution[62] of the difficulty is offered by the use of Wagner's method of earth connexion, which is essentially a means of arranging the potential distribution of the network so that the detector is always brought to earth-potential, and therefore takes up no current in virtue of its earth-capacitance. No point on the four arms of the bridge is directly connected to earth, but the generator is shunted by an impedance, whose time constants are adjusted to be similar to those of the bridge arms. One point on this shunt must then be at the same potential as the corners of the bridge to which the detector is connected (the "detector points") and this point is earth-connected. Equality of potential of the detector points and the earth-connected point is tested by connecting the detector to these points, and the adjustment of the potential distribution is made by balancing the auxiliary impedance in much the same way as for the ordinary bridge balance. The necessity of obtaining simultaneously the condition of balance of both bridge and auxiliary shunt is the chief difficulty of the measurement, since the two sets of adjustment required are not quite independent, and frequently the final condition of simultaneous balance is only slowly approached by a series of alternate adjustments of the two sets of conditions.

Next in importance to the effects of displacement currents are those of the inductances of the various connecting leads. The p.d. arising in any lead on account of its self inductance, or of its mutual inductance with any other lead, is proportional to the frequency, and thus becomes of more and more importance as the frequency increases. All the connecting leads are therefore made short and rigid, so as to make the self inductances small and constant, and they are disposed so as to make the mutual inductances as small as possible. Equal ratio arms are generally employed and the whole structure is symmetrically built so that the stray inductances tend to balance one another. As far as possible all the inductive coils in the system, including the generator and detector, are made of toroidal form, so that their mutual inductance with any other coil is negligibly small.

Generally speaking the simpler the bridge network, the greater its chance of success at radio-frequencies. The condenser is the most satisfactory component for such bridges, since it is easily screened and the effect of the screen on the capacitance value is known. Resistance standards are next in order of convenience, but they must be constructed so as to be as far as possible free from inductance, and the

effects of the capacitance of their shields must be allowed for. Short straight thin wires or narrow loops of wire form satisfactory resistances of low value, but complicated coil windings and resistance boxes with switches should only be used with caution. Inductance standards are less satisfactory for high-frequency bridge work, although they have been used at frequencies of the order of 50,000 c./sec. The effects of the capacitance between shield and coils are however comparatively large and more difficult to determine than those in condensers and resistances. Thus most radio-frequency bridges are simple four-armed networks including only resistances and condensers. Their most successful application has been to the measurement of the capacitance and power factor of condensers and dielectrics. Measurements of power factor to 1×10^{-5} can be made at frequencies up to a million[62]. Large impedances of any kind may be measured with the capacitance bridge by connecting the unknown impedance in parallel with a condenser in one arm of the bridge, and measuring the apparent change of capacitance and power factor (or conductance) produced.

Almost any type of valve-generator may be used to supply current to the bridge, provided the conditions discussed above are complied with, but it is desirable that the output shall be independent of the load, otherwise it is difficult to decide when the condition of balance is being approached. Many forms of detector, including the thermo-galvanometer and the thermionic voltmeter, have been used, but the heterodyne detector is now used almost exclusively for accurate work. An auxiliary generator is required in order to produce a beat note of audible frequency, with the current received from the bridge. This audible note is then easily amplified to almost any desired extent, so that very great sensitivity may be obtained.

§ 8. MAGNETIC MEASUREMENTS

New magnetic materials. On the whole there has been comparatively little advance in the subject of magnetic measurements in the last few years. There has been nothing analogous to the discovery of the thermionic valve, which has so greatly influenced many electrical measurements. A few new magnetic materials have, however, been produced, which have had far-reaching effects in the design of many instruments, and which have made it possible to increase the sensitivity available in some magnetic measurements. The most important of these materials are nickel-iron alloys of very high permeability. All the nickel-iron alloys containing more than 30 per cent of nickel are found to display high permeability in weak magnetic fields. The term "permalloy" was at first applied to the 78·5-per-cent nickel alloy which gave the maximum permeability, but it is now used to refer to the whole range of alloys, the composition being indicated by prefixing the figure giving the percentage nickel content, e.g. 78·5 permalloy. The permeability is greatly dependent on heat treatment, e.g. an annealed specimen giving a maximum permeability of 10,000 was by suitable heat treatment made to give a value exceeding 80,000. The value is also greatly affected by mechanical stress and by the presence of traces of impurities. These very high permeabilities are only obtained for very weak field-

strengths, and are not generally accompanied by a high saturation value of intensity of magnetization. Thus 50 permalloy gives a higher saturation value than 78·5 permalloy, although the latter gives the higher maximum permeability. The hysteresis loss for these materials is very small. For certain applications the addition of a small percentage of various non-magnetic metals, e.g. chromium and molybdenum to permalloy, has been found to confer advantages. Thus the saturation value may be lowered, but the initial permeability greatly increased. The presence of the third constituent also in general reduces the sensitiveness of the alloys to heat treatment and mechanical work. The resistivity also may be considerably increased, and this is of course of the first importance in all alternating current work. By a suitable choice of the proportions of iron and nickel, and the addition of small percentages of other metals, alloys having a wide range of magnetic properties and electrical resistivity have been obtained, and these are on the market under trade names such as mumetal, radiometal, rhometal and permax. Typical applications of these high permeability alloys are the construction of magnetic shields for galvanometers and other electrical instruments, the construction of transformers almost free from power loss for all kinds of instrument work, the increasing of the sensitivity of moving-iron instruments including relays and telephone receivers, and the loading of telephone cables. For telephone work, a very desirable property is constancy of permeability over a considerable range of magnetizing force starting from zero. This means of course that the inductance of a coil wound with the particular material as a core would not vary with the current passing through it, over the working range. It is found that certain alloys of iron nickel and cobalt, called perminvars, possess this property to a remarkable degree. The initial permeability may be very high, and may remain practically constant up to say $H = 2$ oersteds, while the hysteresis loss is almost negligible. With such materials it is possible to make coils possessing large and constant inductances and very small resistances. They are widely used as loading coils in telephone work.

When ferromagnetic materials are employed in apparatus required for high-frequency work, the effective permeability is limited by eddy currents in the material. At low frequencies lamination is sufficient to reduce the eddy currents to permissible proportions, but as the frequency is increased it becomes necessary to use material in a fine state of subdivision, the particles being smaller and more perfectly insulated from one another, the higher the frequency. Thus powdered permalloy has been largely used for the cores of telephone loading coils, while more recently several ferromagnetic composite materials have been produced which are suitable for use at radio frequencies. The overall permeability of the material is of course much reduced by the necessary separation of the ferromagnetic particles, but values of about 10 have been obtained with materials suitable for use at frequencies up to about 10^6 c./sec. By the use of these materials, coils of given inductance and effective resistance can be made of much smaller bulk, and with fewer turns of wire than are necessary for air-core coils. They are therefore specially useful for the construction of toroidal coils, which are often required in laboratory work, on account of their freedom from stray magnetic fields. Such coils are easily screened without

appreciably increasing their effective resistance. Schneider[68] has recently given an account of the construction of "ferrocart," an iron powder material produced in Germany. It consists of thin layers of iron particles separated by sheets of thin paper and compressed. During manufacture the particles are subjected to the action of a magnetic field, which gives them some degree of alignment, and is considered to improve their insulation from one another. Another material of this nature which has been produced in America is described by Polydoroff[69]. It consists of iron particles of diameter about 5 to 10 microns, separated from one another by a synthetic resin varnish, the film being about 1 micron thick. Similar materials have been developed in this country, but no particulars appear to have been published. A general survey of the information available concerning such materials has been given by Colebrook[70].

Other new materials of outstanding importance are the magnet steels. The cobalt-chromium steel of Honda with a coercive force of 300 oersteds was a great advance on its predecessors, and quite recently Mishima[71] has produced a nickel-aluminium steel with a coercive force of 700 oersteds. The great importance of these materials is that they make it possible to produce permanent magnets of a given strength with a smaller weight of material. It therefore becomes possible to increase the sensitivity of all moving-magnet instruments, while all instruments containing fixed permanent magnets may be made lighter and more compact.

The measurement of magnetic fields. A new apparatus for the measurement of the earth's magnetic field has been described by L. F. Bates[72]. The method employed is an electrical one, somewhat similar to those of Schuster and Smith, and Dye. A Helmholtz coil system is arranged to produce a uniform magnetic field at its centre, equal and opposite to that of the component of the earth's field which it is desired to measure. The current through the coils is adjusted until the compensation is perfect, i.e. until the resultant magnetic force in the given direction at the centre of the coils becomes zero. The value of the field is then calculated from the value of the current which is measured, and the linear dimensions and numbers of turns of the coils. The new method differs from its predecessors only in the means adopted to detect the condition of balance. The Schuster-Smith apparatus for the measurement of H used a small permanent magnet suspended by a quartz fibre as in the ordinary moving-magnet galvanometer. Dye's instrument for the measurement of V made use of a small coil suspended by means of phosphor-bronze strip as in the moving-coil vibration-galvanometer, except that the suspension was horizontal instead of vertical. Alternating current was passed through the coil, and the tension of the suspension adjusted until the period of vibration of the coil was equal to that of the current, so that resonance was established, and great sensitivity obtained exactly as in the case of the Campbell unifilar vibration-galvanometer. The Bates instrument also uses a suspended moving coil as detector of compensation. Direct current is passed through the coil and the criterion of compensation is that there shall be no deflection of the coil when the current passing through it is reversed. The sensitivity is increased by providing the suspended coil with a core of material, mumetal, of very high permeability; the coil becomes a small electromagnet, the

polarity of which is reversible. The arrangement can be used for the measurement of either H or V, since the suspension of the moving electromagnet is under tension so that it can be given any desired orientation. The axis of the suspension must of course be set at right angles to the component of force which is to be measured, and in the plane of the magnetic meridian. It will be obvious that the sensitivity can be increased by increasing the moment of the electromagnet and by diminishing the torsional control of the suspension. Thus the magnet must be comparatively long and thin so as to make its moment large for a given total weight.

In many magnetic measurements it is necessary to produce a uniform magnetic field throughout a given volume, and for this purpose the Helmholtz coil system referred to above, and the standard solenoid which must be very long, are widely used. In both cases the overall dimensions of the apparatus are very large compared with those of the region of uniform field obtained. Two arrangements for improving the uniformity of the field inside comparatively short solenoids have recently been described[73]. In one, circular compensating coils are placed near the ends of the uniformly wound solenoid (a principle previously used in various forms of permeameter); in the other a second solenoid of larger diameter than the first is placed coaxially and symmetrically with it, and wound in opposition to it. In either case the auxiliary and main windings are connected in series. The relative dimensions of the coils and number of turns required for maximum uniformity of field are calculated, and it is stated that under the conditions given, a coil 40 cm. long and 6 cm. in diameter will provide a field uniform to ± 7 parts in 100,000 over a region 10 cm. long and of 5 cm. diameter.

The measurement of the magnetic properties of materials. (*a*) *Moving-coil instruments.* The measurement of the magnetic properties of ferro-magnetic materials is in the great majority of cases made by means of ballistic galvanometer methods, but Gerdien and Neumann[74] have recently described a magnetometer giving high sensitivity over a great range of magnetizing forces, and therefore suitable for all the more ordinary measurements. It is a moving-coil instrument, i.e. the deflecting system consists of a small suspended coil carrying a current of a few milliamperes as in the Bates magnetometer mentioned above. This principle had previously been employed by Simon and Madelung. In order to make the instrument insensitive to stray fields, the moving system was made astatic by the use of two similar coils connected in series and wound in opposite directions. They were mounted on the same suspension and rigidly connected, one vertically above the other and 50 cm. distant from it. The advantages claimed for the moving-coil as against the more usual moving-magnet instrument are that the sensitivity of the instrument is easily controlled over a wide range, and that departures from the astatic condition due to temperature differences between the upper and lower moving elements are avoided. The astaticism is made as perfect as possible by applying a correcting field to the upper moving coil by means of two systems of Helmholtz coils mounted respectively on E.-W. and N.-S. axes. The specimen under test is magnetized by a field coil and is placed near to the lower moving coil at a distance which is accurately controlled. The action of the field coil itself is compensated by an auxiliary coil on the opposite

side of the lower moving coil, so that the torque exerted on this coil is produced by the magnetization of the specimen alone. This torque is balanced by means of a toroidal coil carrying an independently controlled current, which is arranged to apply a field to the upper moving coil. Thus a null method of measurement is employed, the actual reading taken being the value of the current in the toroidal coil required in order to bring the deflecting system to zero. The instrument is calibrated by means of a standard coil, the field of which can be accurately calculated from its dimensions. Examples are given of the use of the magnetometer to measure the coercive force of cobalt steel, and the permeability curve and hysteresis loop of soft iron. Agreement within about 1 per cent with results given by ballistic galvanometer methods was obtained in all cases.

Another application of the moving-coil principle has been made by Hermann[75] in an apparatus designed for the testing of sheet iron in the form of a single strip. The specimen is magnetized by a solenoid and the value of B is measured by means of a search coil and ballistic galvanometer in the usual way. H is determined in terms of the deflection of a small flat moving coil, carrying a known current, the plane of the coil being parallel and quite close to the side of the specimen, and its axis of rotation perpendicular to the magnetic field, so that its deflection is proportional to the flux density in air in this locality. The coil consists of ten turns of area of the order of 1 cm^2 wound on a thin sheet of mica, which also carries a galvanometer mirror for the observation of the deflection. The suspension is a single fine wire at either end of the coil. This apparatus possesses the advantage that it may also be used for the measurement of the power loss in the material, when it is subjected to alternating magnetization. For this purpose the moving coil is connected to one of the search coils of the specimen, resistance being added, so that the current in the coil is approximately proportional to the voltage across the search coil, i.e. to the rate of change of flux. The strength of the field in which the coil is suspended is as before equal to H in the specimen. Thus the force acting on the moving coil is at any instant proportional to $H.dB/dt$, and since the periodic time of the moving system is very large compared with that of the alternations of current, a steady deflection is obtained which is proportional to the mean rate of energy dissipation in the material. The corresponding value of the flux-density is measured by connecting another B search coil to a sensitive type of rectifier-voltmeter. The value of B is calculated from the mean voltage observed.

(*b*) *Measurements at high flux-densities.* Apparatus for measurements of permeability, etc., at high flux-densities has recently been described by Sanford and Bennett[76] of the Bureau of Standards, and by J. Sugiura[77] of the Electrotechnical Laboratory, Tokio. The general principles are those which are almost universally used for this class of measurement. A specimen of comparatively small cross-section is arranged to complete the magnetic circuit of a massive yoke, carrying powerful magnetizing coils. The value of B is measured by means of a ballistic galvanometer and search coils embracing the specimen, and that of H is obtained similarly by the use of search coils arranged to enclose only the flux in the air near the surface of the material. The main features of the Bureau of Standards instrument are that a

double yoke is used and that the specimen is clamped between the pole-shoes of the two parts. Thus specimens of various cross-sections are easily accommodated. The search coils are permanent parts of the equipment and are attached to the poles of one yoke. In order that the same B search coil may serve for specimens of various cross-sections, it is wound in two portions, one surrounding the specimen, and the other of equal effective area and in opposition to it encloses only flux in air by the side of the specimen. Thus the flux measured by this coil gives $B-H$, while the other coil gives H. The magnetizing coils are wound on the yokes and do not enclose any portion of the specimen, so that the rise in temperature of the specimen due to the heating effect of the magnetizing current is reduced to a minimum. In the Japanese instrument the specimen is placed in the centre of the magnetizing coil and the yoke merely completes the magnetic circuit. H is measured by an annular search coil, and in order to compensate the reading for leakage flux a series of opposed B search coils distributed along the specimen is connected in series with the H coil. The reading is therefore increased or diminished by an amount which is a function of the variation of flux along the specimen. The extent of this compensation is controlled by shunting the subsidiary coils by means of adjustable resistances.

(*c*) *Measurements on high-permeability alloys.* A convenient and accurate method for the measurement of the coercive force of materials which are magnetically very soft, such as the nickel-iron alloys, has been described by F. Stablein[78]. A small straight specimen is mounted in a non-magnetic holder and placed inside a search-coil (connected to a ballistic galvanometer) which is fixed at the centre of two concentric solenoids. The outer and heavier coil provides the magnetizing force required to saturate the material and the second winding only the very small demagnetizing force necessary to reduce the residual flux to zero. After magnetizing the specimen the magnetizing current is gradually reduced to zero (to eliminate the effects of sparking at the switch contacts) and the demagnetizing current is then adjusted until no deflection of the ballistic galvanometer is produced on withdrawing the specimen by means of the holder. Measurements accurate to a few thousandths of an oersted are claimed.

(*d*) *The measurement of power-loss in magnetic materials.* The measurement of the power dissipated in a given quantity of magnetic material when it is subjected to alternating magnetization of a definite frequency and maximum flux-density is a quantity of considerable technical importance. It may readily be measured as the product of the current in the magnetizing coils, and that portion of the back e.m.f. in these coils due to the magnetization. This back e.m.f. may be separated from any other present by winding a second coil round the material of the same number of turns as the magnetizing coil and distributed along the magnetic circuit in the same way. Evidently the only e.m.f. in this circuit will be that induced by the magnetization, and by applying it to the pressure coil of a wattmeter and passing the magnetizing current through the current coils, a measurement of the power losses (due to hysteresis and eddy currents) involved in the magnetization of the circuit may be made. The loss is assumed to be uniformly distributed throughout the magnetic circuit and the loss per unit mass represents an important property of the material.

It is, of course, necessary that the whole magnetic circuit shall be made of the material under test and, in general, strips of the material are built into the form of a square, with either butt or overlapping joints at the corners. There are, of course, departures from uniformity of flux-density at the corners, which must introduce some error, and a circuit built up of ring stampings is better, but these are seldom available. The actual power to be measured will be proportional to the quantity of material used, and even with a sensitive wattmeter a few kilogrammes of material is usually required. It is not however necessary to use a wattmeter for the measurement: the integral of the product of current and induced voltage may also be obtained from a knowledge of the r.m.s. or maximum value of the current, which may be determined by any of the usual methods, and of the ratio of the values of the induced voltage and the current. This ratio is of the nature of an impedance and may therefore be measured by means of an alternating current bridge. When its magnitude and phase angle are known the power loss is easily calculated. Many well-known bridge circuits could be adapted to make the measurement. Four such were described in a recent paper by Hohle[79], but the simple circuit used by Albert Campbell is probably the most convenient for this purpose. The magnetizing coil is connected in series with a variable non-inductive resistance and a variable mutual inductance. The secondary coil of the mutual inductance is opposed to that of the iron circuit, with a vibration galvanometer or telephone in series to indicate balance, and this combined secondary circuit is tapped across the non-inductive resistance. Thus the mutual inductance measures the quadrature component of the induced e.m.f. and the resistance measures the in-phase component. These bridge circuits are much more sensitive than the wattmeter, and may be used for quite small samples of material. They may also be used for higher frequencies. Since the bridge measures the ratio of voltage to current, this method is insensitive to errors arising from inaccuracy in the setting of the flux-density, both numerator and denominator being affected in a similar way. Hohle emphasizes this point. The presence of harmonics causes some difficulty in the bridge methods, since the bridge can only be balanced for one frequency—the fundamental—and therefore may be said to ignore the harmonics. Integration shows that if either the voltage or the current are purely sinusoidal the harmonics make no contribution to the power loss, and no error results. The circuit is generally arranged so that the voltage is sinusoidal. The method then gives agreement with the wattmeter method. Dannatt[80] has recently taken advantage of the very high sensitivity obtainable by the bridge method to adapt it for making localized measurements of power-loss, i.e. whereas previous methods measure the power-loss in the whole magnetic circuit, Dannatt measures the specific loss in part of the circuit only. Thus he is able to make measurements on a single strip of sheet material, or even on different portions of the same strip. To the portion of material to be tested two search coils are applied, one embracing the material as for the measurement of B, and one a flat coil linking only the flux in the air by the side of the iron, as for the measurement of H. These two search coils are connected in turn to a simple alternating current potentiometer (essentially the same as the circuit described above, except that it is in parallel with the magnetizing

winding instead of in series with it) and the magnitudes of the two e.m.f.'s induced in them are measured and also the difference of phase between them. If we represent the alternating magnetic force and magnetic induction by rotating vectors as is usually done for current and voltage, it will be evident that the mean power dissipated per unit volume of the material will be given by $\dot{B}H\cos\theta/4\pi$, where $\dot{B}$ and H are the r.m.s. values of the two vectors representing the rate of change of induction and magnetizing force respectively, and θ is the phase angle between them. Let n_H, A_H and n_B. A_B denote the number of turns and area of cross-section of the H and B coils respectively. Then the e.m.f. induced in the B coil is of magnitude E_B equal to $n_B A_B \dot{B}$, while that induced in the H coil is E_H and equal to $n_H A_H \dot{H}$, which $= n_H A_H \omega H$ where ω denotes the angular frequency and harmonics are ignored. If δ is the measured phase angle between the two e.m.f.'s then since E_B is in phase with $\dot{B}$, while E_H is in quadrature with H, $\cos\theta = \sin\delta$, and the power dissipated per unit volume of the material is given by

$$\frac{\dot{B}H}{4\pi}\cos\theta = \frac{E_H E_B \sin\delta}{4\pi\omega n_H A_B n_B A_B}.$$

The above expressions assume that c.g.s. electromagnetic units are used throughout, but as E_H and E_B will be measured in volts and A_H and A_B in cm^2, the final result must be multiplied by 10^9 in order to obtain the value in watts per cm^3.

The following conditions must be fulfilled for the successful working of the method. The flux in the iron should be purely sinusoidal, so that harmonics in H will introduce no error; the magnetization of the specimen should be uniform, and the value of H across the section of the magnetizing coils should also be uniform. These conditions are realized as follows. A magnetic circuit is formed from the specimen strip, and a similar but longer one with its ends bent round to serve as connecting pieces, lap joints being used. Uniformly distributed magnetizing coils are wound in a number of sections, as low in resistance as possible, and they are all connected in parallel across a sinusoidal voltage supply. Such an arrangement tends to produce constant sine wave flux in all parts of the circuit. The iron strips and search coils are placed between two pieces of copper sheet $\frac{1}{16}$ in. thick parallel to the H search coils and the specimen. Any radial component of H is greatly reduced by the eddy currents set up in the copper sheets, and thus the uniformity of H over the cross-section is improved. Since the B search coil does not surround the copper sheets, the losses in these sheets are not measured.

The measurement of the power losses in a single sheet of material may also be effected by means of the moving-coil magnetometer of Hermann described in a previous paragraph. Dannatt's method is likely to be the more sensitive and accurate, although several of his refinements could equally well be applied to Hermann's or any other method.

Dannatt points out that although the maximum value of B may be deduced from the r.m.s. value of the voltage induced in the B search coil, provided the wave-form of the flux is known to be sinusoidal, it is generally more satisfactory to measure the mean induced voltage by the use of a synchronous rectifying commutator and high

resistance galvanometer. It is not then necessary to make any assumption as to form factor. The same point is discussed in two other recent papers[81,82].

(*e*) *Measurements on feebly magnetic materials.* One of the most interesting investigations of recent years is that of Kapitza and Webster[83] on the measurement of the susceptibilities of materials which are only very feebly magnetic, such as ordinary liquids and gases. The method at present most widely used for this purpose consists in measuring the mechanical force experienced by a sample of the material when it is placed in the field of a powerful magnet. As is well known, the force is zero if the field is uniform, but when it is non-uniform there is a force tending to move paramagnetic bodies towards the strongest part of the field, and diamagnetic bodies towards the weakest, its magnitude being given by the formula

$$F_x = \chi \int_0^V H \frac{dH}{dx}\, dV \qquad \text{.....(5),}$$

where χ is the volume susceptibility, H the magnetic force, and dH/dx the gradient of the magnetic force in the direction x in which the force is measured, in the position occupied by a small element of volume dV of the body. In order that the force may be measurable the body must be delicately suspended, and a torsion balance is generally employed for this purpose. This makes the method almost useless for investigating the changes in magnetic properties produced by changes of physical conditions, such as the application of pressure or an increase of temperature. For such measurements it is essential that the body shall be fixed. The force may still be measured if the magnet may be freely suspended, since the reaction on the magnet must equal the force on the body, but a magnet which is small enough to be suspended from a torsion balance can obviously not be expected to produce a sufficiently powerful field to secure the maximum sensitivity, since the force is proportional to H as well as to dH/dx. The difficulty has been overcome by placing the small suspended magnet and the body under test, which is fixed in a uniform magnetic field of great intensity provided by a large magnet which is also fixed. The suspended magnet consists simply of a small iron rod which becomes magnetized by induction as soon as it is placed in the strong magnetic field. The uniform magnetic field exerts no resultant force on either the body or the iron rod, but as a result of the non-uniform field of the magnetized rod there is a force between the rod and the body, which may be represented by equation (5), and since H now includes the value of the uniform magnetic field, the sensitivity is correspondingly increased. The chief difficulty lies in the fact that it is impossible to produce a very strong magnetic field which is perfectly uniform, and a very slight non-uniformity must cause a force on the suspended rod which will be very large compared with that due to the body under test, owing to the high susceptibility of the rod. On this account two exactly similar iron rods are used, and they are symmetrically mounted on the two arms of a torsion balance. The whole arrangement is then made as far as possible symmetrical about the axis of rotation of the balance arm, so that the forces on the two rods are always equal and such as to cause no deflection of the balance arms. The body under test is placed near one of the rods only and the force it exerts

tends to deflect the balance. A null method of measurement is employed, the zero position of the balance being restored by the force exerted on the rod by the current in a small coil suitably placed near it. The magnetic field is produced by a large Weiss electromagnet (excitation 6 kW.) which will give a value of H of 12,000 oersteds in an air-gap 3·5 cm. wide, between pole faces of diameter 18 cm. The whole balance is quite small. The iron rods, 16 mm. long and 2 mm. diameter, are placed one on either side of the beam, their distance apart being 3 cm. The suspension consists of phosphor bronze strip with a free length of 2 mm. on each side of the beam. The rods and suspension strip are both parallel to the field of the large magnet and the axis of the suspension coincides with the line joining the centres of the pole faces. All the metal parts are made of cadmium-copper which has been found to be the only metal sufficiently non-magnetic for the purpose. Coils through which current can be circulated are placed near each of the iron rods, and these are used for controlling the zero position and the sensitivity of the instrument. A sensitivity corresponding to a deflection of 50 cm. for oxygen at atmospheric pressure can be obtained, but zero instability is the limiting factor. Although special means are adopted for the regulation of the voltage applied to the large magnet, variations of the field with time are always appreciable, their effect being to make the zero position of the balance unstable. Electromagnetic damping is employed, and made adjustable so that the critical condition can be used. It will be evident that the balance reading must be very sensitive to the shape of the test body, and its exact position with reference to the iron rod. The method is therefore only suitable for relative measurements. Liquids and gases are tested by clamping a glass vessel in a definite position near one of the rods and making a measurement with the vessel full and empty. Powders may be similarly treated. Solid bodies cut to the same size and shape may be directly compared, and solid bodies which cannot be cut to size may be investigated by an application of Archimedes' principle. A measurement is made on the body in air, and then one with it immersed in a liquid of known susceptibility. The difference between the two measurements gives the effect of the displaced liquid, which has the same size and shape as the solid body.

A very simple form of torsion balance for the inspection of steels which are comparatively non-magnetic has been described by Sanford[84]. A small magnet pivoted like a dip-needle, but with a torsion head and spiral spring control, is mounted on the horizontal surface of the specimen. One pole of the magnet is brought into contact with the surface, and the torsion which must then be applied in order to break the contact is measured. In order to eliminate the effect of the earth's field on the magnet, the system is made astatic. The pivoted needle is formed from two similar magnets of cobalt steel, which are mounted end to end, with like poles in contact. It is found that the magnetic properties of the austenitic steels are closely related to their power of resisting corrosion, so that a rapid test of this kind is of considerable value.

A simple torsion balance has also been used by Otis[85] for measurements on ferromagnetic material in the form of very thin films. A disc with the film on it is suspended in a glass case from a torsion head, and the whole inserted in the field

of a solenoid. The torsion which must be applied in order to deflect the specimen through a given angle is measured, and gives a measure of the magnetic induction in the specimen.

Probably the most generally convenient simple test for non-magnetic bodies consists in taking two coils of the largest available mutual inductance and passing a constant current through one, while the other is connected to a sensitive ballistic galvanometer. The body to be tested is inserted in the field of the coils and then withdrawn. The throw of the galvanometer on insertion or withdrawal gives a measure of the susceptibility of the specimen. Zero-instability due to fluctuation of current may be avoided by the use of two similar mutual inductances in opposition, the body being inserted in the field of one only. The method has long been known and widely used, but G. Grenet[86] has recently called attention to its usefulness for work on rocks.

Miscellaneous applications of magnetic methods of measurement. Several arrangements[87, 88] have been described for the application of magnetic measurements to the measurement of mechanical force. The magnetic properties of materials like permalloy are very sensitive to the application of mechanical stress: the permeability may change by 20 per cent on the application of a load of one-tenth of the breaking load (the corresponding value for ordinary steel may be only 1 per cent). Thus if a suitable specimen of permalloy is wound with a magnetizing coil, and a bridge circuit is then arranged to give an indication of the inductance of the coil, the bridge reading will also be a function of the applied force, and the arrangement may be calibrated so as to give readings of mechanical force directly.

Wire ropes[89, 90] are conveniently tested for flaws by a magnetic method. The rope s drawn continuously through two coils, one which magnetizes it to saturation and one which acts as a search coil, and is connected to a recording galvanometer, which records the existence of any discontinuities. By magnetizing to saturation, the effects of minor irregularities of no importance, such as changes in permeability due to slight variations in heat treatment, etc., are smoothed over. The magnetic testing of welds has also been investigated but so far the method has not been conspicuously successful.

A magnetic method of measuring thickness[91] has also been described, but it is of very limited application. The magnetic circuit of a transformer is constructed in two parts in parallel exactly similar to one another and each containing an air-gap. A standard specimen is inserted in one gap, and the plate under test in the other. The secondary coils on the two parts of the circuit are connected to a differential a.-c. galvanometer, which is calibrated to give a direct reading of thickness.

§ 9. UNITS AND STANDARDS

At the present time the great bulk of electrical experimental work is recorded in terms of the eight internationally recognized practical units, viz. the ampere, ohm, volt, watt, joule, coulomb, farad and henry. The international ampere was defined in 1908 as the current which deposits silver at a certain standard rate, and the inter-

national ohm as the resistance of a column of mercury of a certain length and mass. The other units are derived from these two, and thus they have been arbitrarily defined in terms of the properties of silver and mercury. This procedure was regarded as undesirable by many physicists even in 1908, and the discovery of isotopes has done something further towards shaking its foundations. Nevertheless the units have been almost universally used in both engineering and physics. The physicist usually regarded Maxwell's c.g.s. electromagnetic system of units as his real basis, and regarded the practical units merely as convenient decimal multiples of these. The practical units were of course chosen so that this was approximately true; the discrepancy is at most of the order of 0·05 per cent, which gives sufficient accuracy for ordinary purposes. The accuracy of measurements of all kinds is, however, steadily improving, and the fact that, for example, the accepted unit of electrical power, the international watt, is inconsistent with the corresponding mechanical unit must be regarded as a serious objection to the present international units. The work of the chief national standardizing laboratories has shown that electrical quantities can now be measured in terms of the mechanical units and the electrical and magnetic properties of free space (in other words absolute electrical measurements can be made) with an accuracy approaching 2 parts in 10^5. The international units are not reproducible to a higher degree of accuracy than this, so that there is no further justification for their use, and it is now proposed[92] to abandon them and to adopt an "absolute" system, i.e. a system derived from the mechanical units and independent of the electrical properties of any material. The decision was reached at a recent General Conference on Weights and Measures, but the announcement of the numerical relation between the new units and the old has been deferred until 1935 when, it is hoped, absolute measurements now in progress in several countries will be completed. Arising partly from this decision there have been many recent papers[93] on systems of electrical and magnetic units, and on nomenclature. The International Committee recommended the adoption of the "absolute system," but it is to be remembered that there are many absolute systems, and even if it is assumed that the eight practical units are to be used after only very small adjustments have been made in their value, there remains more than one possibility. It is entirely a question as to what are to be regarded as the fundamental units. By choosing the centimetre-gramme-second, and the permeability of free space, Maxwell derived his electromagnetic system, and this has frequently been regarded as "the absolute system," and it has been assumed that the new units would be obtained from those of this system by multiplying by a power of ten. Thus the new ohm would differ from the old by about 1 part in 2000. This is probably the course contemplated by the International Committee, but G. A. Campbell[93, 94] has suggested that in order to avoid the necessity of altering the values of precision resistances the ohm should be taken as the fourth fundamental unit. He points out that the units of length, mass, and time are purely arbitrary, and that the fourth may as well be arbitrary. It could therefore be given the value at present in general use. The chief difficulty is that of preserving the value. This of course would mean that the permeability of free space would no longer be unity but a fundamental

constant to be determined and recorded, just as that equally fundamental constant, the velocity of light, is recorded.

The suggestion of adopting an absolute system of units instead of the present systems has revived the question whether it is not possible to make the proposed system complete, so that it may include magnetic as well as electrical units, and to choose the system so that it may possess all the advantages of the various rival systems now in use, while the rest will serve no useful purpose and will gradually disappear. At the present time the experimental physicist appears to prefer practical units for all electrical quantities, but he uses c.g.s. electromagnetic units for magnetic quantities, presumably because magnetic units were not included in the practical series. Thus at a meeting of the International Electrotechnical Commission at Oslo in 1930 international agreement was reached on the use of the names gilbert, oersted, maxwell and gauss for the c.g.s. electromagnetic units of magnetomotive force, magnetic intensity, flux and magnetic induction respectively, whereas the electrical units of the system have not been named, since they are seldom used. The electric moments of molecules are often recorded in c.g.s. electrostatic units, and this system is frequently employed in problems which are mainly governed by the laws of electrostatics. Many mathematical physicists on the other hand prefer to adopt the Heaviside, Lorentz, or Gaussian systems in which the form of the fundamental electromagnetic equations is such as to show symmetry between electric quantities and the corresponding magnetic quantities. Leigh Page[93] has proposed a symmetrical system of units including the eight practical units already internationally recognized. Kenelly has also suggested a symmetrical system. These systems would however include units which are at present unfamiliar. G. A. Campbell advocates a system which combines most of the advantages of the others, and at the same time uses the units most widely used at the present time. The fundamental units are the metre, kilogram, second, and ohm; the unit of magnetomotive force is the ampere-turn; and that of magnetic intensity the ampere-turn per metre. As Campbell points out, much the same proposal has been made by Giorgi and by Robertson and it has been supported by engineers in several countries, particularly with a view to avoiding the use of 4π in problems which do not involve a spherical distribution. Moreover since the system is complete and absolute, it would meet the requirements of the physicist, who already uses most of the units.

In connexion with formulation of a preferred system of units there has been much discussion[95,96,97] of the question whether magnetic permeability is to be regarded as a quantity having dimensions or a mere number. Abraham[93] insists that the dimensions of physical quantities are matters of arbitrary convention and that no further significance should be attached to dimensions, but among engineers there is a strong desire to regard the relation between magnetic force and magnetic induction as analogous to that between mechanical force and displacement and between electric force and displacement. Magnetic force and induction are therefore regarded as quantities of a different physical character, and a convention which assigns to them different dimensions is an aid to thought, and is preferred. Thus in the system of units proposed by G. A. Campbell, permeability would be expressed in a unit which

is at present unfamiliar, but magnetic practice would be mainly concerned with specific permeability, i.e. the ratio of the permeability of the material in question to that of free space, and this of course would be a mere number identical in value with the existing recorded value of the permeability of the material.

In order to simplify as far as possible the equations linking together all the various units of the complete system, it has been suggested[98] that electrical quantity should replace mass as one of the fundamental units. G. A. Campbell[93] points out that the metre, second, watt and coulomb would be very convenient in this respect.

§ 10. SCOPE OF THIS REPORT

In this report the more conspicuous of the developments in electric and magnetic measurements that have occurred in 1933, and to some extent in the preceding few years, have been reviewed. Many lines of development have been ignored in the opinion that they could be dealt with to better advantage in future reports.

§ 11. ACKNOWLEDGMENTS

The author wishes to acknowledge his indebtedness to his colleagues, particularly Messrs C. E. Webb, L. H. Ford, W. H. Ward, and L. Essen, for assistance in surveying the literature.

REFERENCES

(1) G. Ising. *Phil. Mag.* **1**, 827 (1926).
(2) L. S. Orstein, H. C. Burger, J. Taylor and W. Clarkson. *Proc. roy. Soc.* A, **115**, 391 (1927).
(3) Moll and Burger. *J. sci. Instrum.* **10**, 223 (1933).
(4) G. Ising. *Ann. Phys.*, Lpz., **14**, 755 (1932).
(5) L. Meylan. *Helv. phys. Acta*, **5**, 276 (1932).
(6) H. Sell. *Z. tech. Phys.* **13**, 320 (1932).
(7) W. B. Nottingham. *J. Franklin Inst.* **209**, 348 (1930).
(8) C. E. Wynn-Williams. *Proc. Camb. phil. Soc.* **23**, 810 (1927).
(9) —— *Phil. Mag.* **6**, 324 (1928).
(10) A. S. MacFarlane. *J. sci. Instrum.* **10**, 142 (1933).
(11) —— *Phil. Mag.* **14**, 1 (1932).
(12) G. F. Metcalf and B. J. Thompson. *Phys. Rev.* **36**, 1489 (1930).
(13) L. A. Dubridge. *Phys. Rev.* **37**, 392 (1931).
(14) W. Soller. *Rev. sci. Instrum.* **3**, 416 (1932).
(15) J. A. C. Teegan and A. M. Hayes. *J. sci. Instrum.* **10**, 110 (1933).
(16) P. A. MacDonald and E. M. Campbell. *Physics*, **4**, 237 (1933).
(17) P. A. MacDonald and J. T. MacPherson. *Phil. Mag.* **15**, 72 (1933).
(18) P. Donzelot and J. Divoux. *C.R. Acad. Sci.*, Paris, **196**, 1579 (1933).
(19) S. B. Ellis and S. J. Kiehl. *Rev. sci. Instrum.* **4**, 131 (1933).
(20) R. E. Burroughs and J. E. Ferguson. *Rev. sci. Instrum.* **4**, 406 (1933).
(21) J. Roulleau. *C.R. Acad. Sci.*, Paris, **196**, 1786 (1933).
(22) G. Keinath. *Physik*, **1**, 21 (1933).
(23) H. Pfannenmüller. *Arch. tech. Mess.* p. 540 (1932).
(24) L. O. Grondahl. *Rev. Mod. Phys.* **5**, 141 (1933).
(25) C. H. Walter. *Z. tech. Phys.* **13**, 363 (1932).

(26) K. Kreielsheimer. *Ann. Phys.*, Lpz., **17**, 293 (1933).
(27) W. Meissner and V. Adelsberger. *Z. tech. Phys.* **14**, 111 (1933).
(28) —"— *Z. tech. Phys.* **13**, 475 (1932).
(29) —— —— *Z. tech. Phys.* **11**, 102 (1930).
(30) E. Giebe and A. Scheibe. *Hochfrequenztech. u. Elektroakust.* **41**, 83 (1933).
(31) —— —— *Jb. drahtl. Telegr.* **35**, 165 (1930).
(32) Heinrich Hertz Institute. *Elekt. Nachr.-Tech.* **10**, 229 (1933).
(33) G. W. Fox and M. Underwood. *Physics*, **4**, 10 (1933).
(34) S. Matsumura and S. Ishikawa. *Rep. Radio Res. Japan*, **3**, 1 (1933).
(35) V. Petržilka. *Ann. Phys.*, Lpz., **15**, 72 (1932).
(36) V. J. Andrew. *Proc. Inst. Radio Engrs*, N.Y., **21**, 982 (1933).
(37) C. E. Smith. *Proc. Inst. Radio Engrs*, N.Y., **21**, 37 (1933).
(38) A. Scheibe and V. Adelsberger. *Ann. Phys.*, Lpz., **18**, 1 (1933).
(39) H. Kono. *Rep. Radio Res. Japan*, **3**, 127 (1933).
(40) R. Yoneda and K. Yamaguchi. *Res. electrotech. Lab. Minist. Commun. Japan*, No. 355 (1933).
(41) H. L. Curtis and C. Moon. *Bull. U.S. Bur. Stand.* **22**, 487 (1927).
(42) Gruneisen and Giebe. *Wiss. Abh. phys.-tech. Reichsanst.*, Berl., **5**, 81 (1922).
(43) H. L. Curtis, C. M. Sparks, L. Hartshorn and N. F. Astbury. *Coll. Res. nat. phys. Lab.*, **24**, 57 (1932).
(44) —— —— —— —— *Bur. Stand. J. Res.* **8**, 507 (1932).
(45) E. Giebe and G. Zickner. *Z. InstrumKde*, **53**, 1 (1933).
(46) B. G. Churcher and C. Dannatt. *J. Instn. elect. Engrs*, **69**, 1019 (1931).
(47) H. Lampe. *Z. Phys.*, Lpz., **79**, 254 (1932).
(48) J. Clay. *Z. Phys.*, Lpz., **78**, 250 (1932).
(49) A. C. Bartlett. *J. sci. Instrum.* **8**, 260 (1931).
(50) Drake, Pierce and Dow. *Phys. Rev.* **35**, 613 (1930).
(51) J. G. Malone, L. O. Case and A. L. Ferguson. *J. chem. Phys.* **1**, 842 (1933).
(52) M. Seeberger. *Ann. Phys.*, Lpz., **16**, 77 (1933).
(53) K. Kohl. *Elektrotech. Nachr.* **6**, 354 (1929).
(54) E. Glowatzki. *Ann. Phys.*, Lpz., **18**, 217 (1933).
(55) S. Sugden. *J. chem. Soc.* p. 768 (1933).
(56) F. M. Colebrook. *Wireless Engr*, **8**, 639 (1931).
(57) L. Rohde and W. Schlegelmilch. *Elektrotech. Z.* **24**, 580 (1933).
(58) H. Barkhausen and K. Kurz. *Phys. Z.* **21**, 1 (1920).
(59) E. C. S. Megaw. *J. Instn. elect. Engrs*, **72**, 313 (1933).
(60) H. N. Kozanowski. *Proc. Inst. Radio Engrs*, N.Y., **20**, 957 (1932).
(61) E. C. S. Megaw. *J. Instn. elect. Engrs*, **72**, 326 (1933).
(62) D. W. Dye and T. I. Jones. *J. Instn. elect. Engrs*, **72**, 169 (1933).
(63) L. Behr and A. J. Williams. *Proc. Inst. Radio Engrs*, N.Y., **20**, 969 (1932).
(64) K. Schlesinger. *Elektrotech. Z.* **52**, 533 (1931).
(65) W. J. Shackleton and J. G. Ferguson. *Bell Syst. tech. J.* **6**, 142 (1927).
(66) —— —— *Bell Syst. tech. J.* **7**, 70 (1928).
(67) —— —— *Bell Syst. tech. J.* **8**, 560 (1929).
(68) A. Schneider. *Wireless Engr.*, **10**, 183 (1933).
(69) W. J. Polydoroff. *Proc. Inst. Radio Engrs*, N.Y., **21**, 690 (1933).
(70) F. M. Colebrook. *Magnetic Materials at Radio Frequencies.* D.S.I.R. Radio Research. Special Report, No. 14 (H.M. Stationery Office.)
(71) Mishima. *Ohm*, **19**, No. 7 (1932).
(72) L. F. Bates. *Proc. phys. Soc.* **45**, 181 (1933).
(73) A. Bühl and F. Coetener. *Phys. Z.* **33**, 773 (1932).
(74) H. Gerdien and H. Neumann. *Wiss. Veröff. Siemens-Konz.* **11**, No. 2, 12 (1932).
(75) P. C. Hermann. *Z. tech. Phys.* **13**, No. 11, 541 (1932).
(76) R. L. Sanford and E. G. Bennett. *Bull. U.S. Bur. Stand.* **10**, 567 (1933).
(77) J. Sugiura. *Res. electrotech. Lab. Minist. Commun. Japan*, No. 354 (1933).
(78) F. Stablein. *Z. tech. Phys.* **13**, No. 11, 532 (1932).

(79) W. Hohle. *Arch. Elektrotech.* **25**, 813 (1931).
(80) C. Dannatt. *J. sci. Instrum.* **10**, 276 (1933).
(81) B. M. Smith and C. Concordia. *Elect. Engng*, **51**, 36 (1932).
(82) E. A. Neumann and J. Pfaffenberger. *Arch. Elektrotech.* **27**, 287 (1933).
(83) P. Kapitza and W. L. Webster. *Proc. roy. Soc.* **132**, 442 (1931).
(84) R. L. Sanford. *Bull. U.S. Bur. Stand.* **10**, 321 (1933).
(85) H. N. Otis. *Rev. sci. Instrum.* **4**, 681 (1933).
(86) G. Grenet. *C.R. Acad. Sci.*, Paris, **196**, No. 12, 874 (1933).
(87) G. Keinath. *Arch. tech. Mess.* **132**, 1 (1932).
(88) G. Sacerdote. *Electrotechnics*, **19**, 485 (1932).
(89) H. Herbst. *Z. Ver. dtsch. Ing.* **76**, 557 (1932).
(90) F. Wever and A. Otto. *Mitt. K.-Wilh.-Inst. Eisenforsch.* **12**, 373 (1930).
(91) P. Bricot. *C.R. Acad. Sci.*, Paris, **196**, No. 10, 689 (1933).
(92) *Procès-Verbaux de Comité Consultatif d'Electricité et de Photométrie* (1933).
(93) Sir Richard Glazebrook, H. Abraham, Leigh Page, G. A. Campbell, H. L. Curtis and A. E. Kenelly. "Systems of Electrical and Magnetic Units." *Bull. nat. Res. Coun.* No. 93 (1933).
(94) G. A. Campbell. *Physics*, **3**, 230 (1932).
(95) R. T. Glazebrook. *J. Instn. elect. Engrs*, **72**, 265 (1933).
(96) A. E. Kenelly. *Proc. Amer. phil. Soc.* **72**, 39 (1933).
(97) E. Griffiths. *Nature*, Lond., **130**, 987 (1932).
(98) W. Cramp. *Nature*, Lond., **130**, 368 (1932).

CAMBRIDGE: PRINTED BY WALTER LEWIS, M.A., AT THE UNIVERSITY PRESS